

=> d his

FILE 'HCAPLUS' ENTERED AT 15:14:38 ON 29 DEC 1999  
 E LIPPS B/AU  
 L1 9 S E5,E9  
     E LIPPS F/AU  
 L2 6 S E3,E4,E6  
 L3 11 S L1-L2  
 L4 4 S LTNF  
 L5 29 S LETHAL? (L) TOXIN? (L) NEUTRALI? (L) FACTOR

**Point of Contact:**  
**Jan Delaval**  
**Librarian-Physical Sciences**  
**CM1 1E01 Tel: 308-4498**

FILE 'REGISTRY' ENTERED AT 15:17:52 ON 29 DEC 1999  
 L6 10 S E1-E10  
 L7 4 S L6 AND SQL/FA  
 L8 3 S L7 NOT MAN/CI  
     E LKAMD/SQEP  
 L9 1 S E3  
     E LKAMDPTPPL/SQEP  
 L10 1 S E3  
     E LKAMDPTPPLWIKTE/SQEP  
 L11 1 S E3  
 L12 3 S L9-L11

FILE 'HCAPLUS' ENTERED AT 15:21:07 ON 29 DEC 1999  
 L13 2 S L12  
 L14 3 S L3 AND L4,L5  
 L15 2 S L3 AND L13  
 L16 3 S L13-L15  
 L17 3 S LETHAL? TOXIN? NEUTRAL? FACTOR  
 L18 0 S ANTILETHAL? TOXIN? NEUTRAL? FACTOR  
 L19 0 S ANTI TNF  
 L20 3 S L16,L17

FILE 'BIOSIS' ENTERED AT 15:25:26 ON 29 DEC 1999  
 E LIPPS B/AU  
 L21 9 S E3,E6,E7  
     E LIPPS F/AU  
 L22 0 S LTNF  
 L23 0 S L17  
 L24 0 S L12  
 L25 0 S L21 AND ?TOXIN?  
 L26 0 S L21 AND OPPOSUM

FILE 'WPIDS' ENTERED AT 15:27:05 ON 29 DEC 1999  
 L27 0 S LNTF OR ANTI LNTF  
 L28 2 S LETHAL? TOXIN? NEUTRAL? FACTOR  
     E LIPPS B/AU  
 L29 5 S E5,E8  
 L30 5 S L28,L29

=> fil reg  
 FILE 'REGISTRY' ENTERED AT 15:32:32 ON 29 DEC 1999  
 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
 PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
 COPYRIGHT (C) 1999 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 28 DEC 99 HIGHEST RN 251933-56-9  
 DICTIONARY FILE UPDATES: 28 DEC 99 HIGHEST RN 251933-56-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 13, 1999

Please note that search-term pricing does apply when  
 conducting SmartSELECT searches.

POTENTIAL STEREO BOND SEARCH PROBLEM WITH STN EXPRESS WITH DISCOVER!

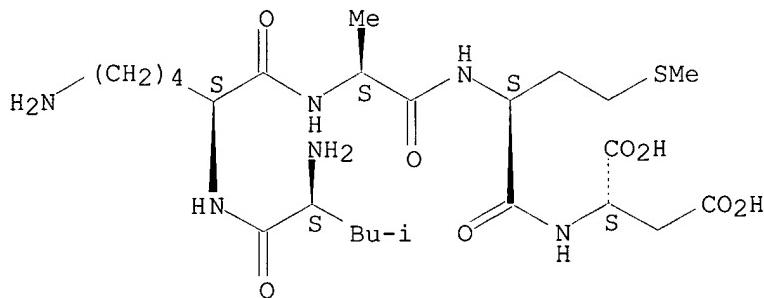
5.0 (Windows Only) SEE NEWS MESSAGE FOR DETAILS.

=> d sqide can tot l12

L12 ANSWER 1 OF 3 REGISTRY COPYRIGHT 1999 ACS  
 RN 206755-64-8 REGISTRY  
 CN L-Aspartic acid, L-leucyl-L-lysyl-L-alanyl-L-methionyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 SQL 5

SEQ 1 LKAMD  
 =====  
 HITS AT: 1-5  
 MF C24 H44 N6 O8 S  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1967 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

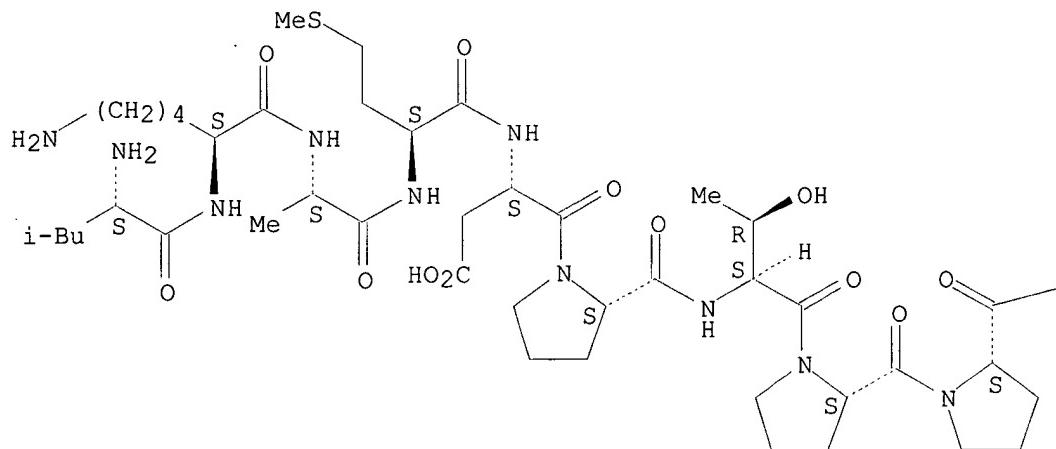
REFERENCE 1: 128:305070

L12 ANSWER 2 OF 3 REGISTRY COPYRIGHT 1999 ACS  
 RN 206755-63-7 REGISTRY  
 CN L-Leucine, L-leucyl-L-lysyl-L-alanyl-L-methionyl-L-.alpha.-aspartyl-L-prolyl-L-threonyl-L-prolyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 SQL 10

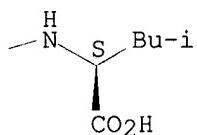
SEQ 1 LKAMDPTPPL  
 =====  
 HITS AT: 1-10  
 MF C49 H83 N11 O14 S  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1967 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

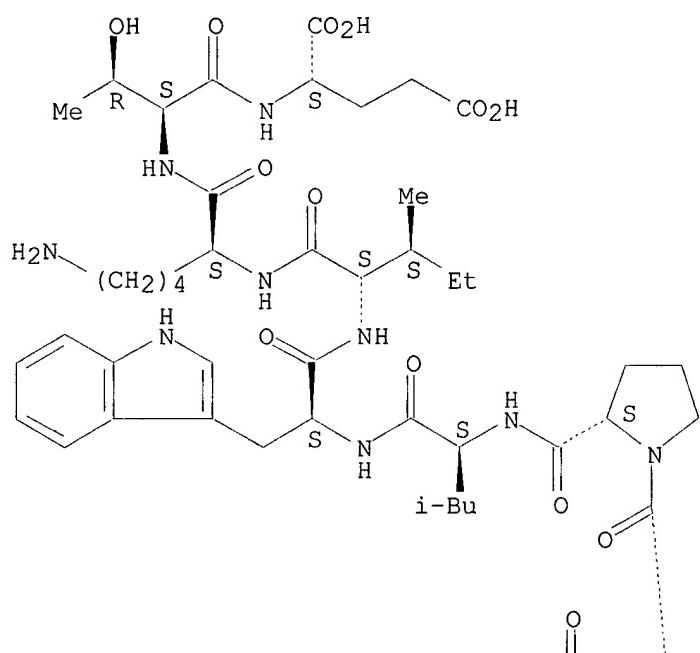
REFERENCE 1: 128:305070

L12 ANSWER 3 OF 3 REGISTRY COPYRIGHT 1999 ACS  
 RN 184782-42-1 REGISTRY  
 CN L-Glutamic acid, L-leucyl-L-lysyl-L-alanyl-L-methionyl-L-.alpha.-aspartyl-L-prolyl-L-threonyl-L-prolyl-L-prolyl-L-leucyl-L-tryptophyl-L-isoleucyl-L-lysyl-L-threonyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 SQL 15

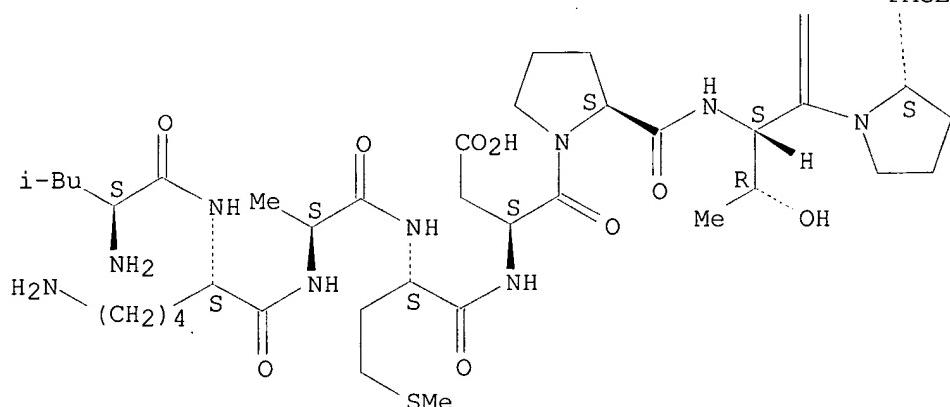
SEQ 1 LKAMDPTPPL WIKTE  
 ====== ======  
 HITS AT: 1-15  
 MF C81 H130 N18 O22 S  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXLIT, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



2 REFERENCES IN FILE CA (1967 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 128:305070

REFERENCE 2: 126:27960

=> fil hcplus  
 FILE 'HCPLUS' ENTERED AT 15:32:48 ON 29 DEC 1999  
 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
 PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
 COPYRIGHT (C) 1999 AMERICAN CHEMICAL SOCIETY (ACS)

26, 1996), unless otherwise indicated in the original publications.

FILE COVERS 1967 - 29 Dec 1999 VOL 132 ISS 1  
FILE LAST UPDATED: 28 Dec 1999 (19991228/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

This file supports REG1stRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

=> d all tot 120

L20 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 1999 ACS  
AN 1999:268585 HCAPLUS  
DN 131:112584  
TI Anti-lethal factor from opossum serum is a potent antidote for animal, plant and bacterial toxins  
AU **Lipps, B. V.**  
CS Ophidia Products, Inc., Houston, TX, USA  
SO J. Venomous Anim. Toxins (1999), 5(1), 56-66  
CODEN: JVTOFG; ISSN: 0104-7930  
PB Center for the Study of Venoms and Venomous Animals  
DT Journal; (computer magnetic disk)  
LA English  
CC 4-5 (Toxicology)  
AB Currently, the use of antivenoms is the only available treatment for envenomation caused by venomous animals namely, snake, scorpion, spider, tick and jelly fish. Antivenoms are generally produced in large animals, mostly in horses. A large percentage of the population is allergic to horse proteins. Several animals are known to be resistant to snakebites and the antihemorrhagic and anti-lethal components have been isolated from sera of opossum, mongoose, meerkat and hedgehog, as well as from venomous and non-venomous snakes. **Anti-lethal factor** named **lethal toxin neutralizing factor** (**LTNF**) has been isolated in purity from opossum (*Didelphis virginiana*) serum by high pressure liq. chromatog. (HPLC). The mol. wt. of **LTNF** is 63 kDa, and it does not form pptn. with venoms or **toxins** by immunodiffusion. Death due to i.p. (IP) injection of a predetd. LD of venom from major families of snakes, for instance Crotalidae, Elapidae, Viperidae and Hydrophiidae, is prevented in mice by subsequent IP inoculation of **LTNF**. Furthermore, **LTNF neutralizes** the **lethality** of scorpion and bee venoms and **toxins** from various animals, plants and bacteria. Thus, natural **LTNF** from opossum serum has potential as a universal therapy for envenomation caused by animals, plants and bacteria.  
ST opossum blood lethal factor antidote toxin; *Didelphis* blood lethal factor antidote toxin  
IT Antidotes  
Didelphis virginiana  
Poisoning (biological)  
Serum (blood)  
Snake venoms  
Venoms  
(antilethal factor from opossum serum is potent antidote for animal, plant and bacterial toxins)  
IT Ricins  
Toxins  
RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
(antilethal factor from opossum serum is potent antidote for animal, plant and bacterial toxins)  
IT Antitoxins  
RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(antilethal factor from opossum serum is potent antidote for animal,

plant and bacterial toxins)

IT 9007-40-3, Crotoxin 11029-72-4, Holothurin 12584-83-7, Cobratoxin  
52019-39-3, Taipoxin 107231-12-9, Botulin

RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
(antilethal factor from opossum serum is potent antidote for animal,  
plant and bacterial toxins)

IT 9001-84-7, Phospholipase a2

RL: BAC (Biological activity or effector, except adverse); BIOL  
(Biological study)  
(antilethal factor from opossum serum is potent antidote for animal,  
plant and bacterial toxins)

L20 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 1999 ACS  
AN 1998:263232 HCAPLUS  
DN 128:305070  
TI Lethal toxin neutralizing factors  
IN Lipps, Binie V.; Lipps, Frederick W.  
PA USA  
SO U.S., 11 pp. Cont.-in-part of U.S. 5,576,297.  
CODEN: USXXAM  
DT Patent  
LA English  
IC ICM A61K038-10  
      ICS A61K038-16; C07K007-08; C07K014-47  
NCL 514014000  
CC 4-5 (Toxicology)  
FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5744449	A	19980428	US 1996-657163	19960603
	US 5576297	A	19961119	US 1994-310340	19940922
PRAI	US 1993-58387	19930510			
	US 1994-310340	19940922			

AB Opossum whole serum exhibits a life saving property by neutralizing the lethality of venoms from all major families of poisonous snakes, and therefore an injection of Opossum serum can used as a novel treatment for many types of envenomation. Preferably, the injectable treatment for envenomation should be a compn. obtained from the fraction of Opossum whole serum which contains the **LTNF**, i. e. the so called "**LTNF-n**", in purity. A method is given for the manuf. of a **LTNF** from the serum of an opossum (*Didelphis virginiana*) serum, by fractionating the opossum serum and isolating this select fraction from the plurality of fractions having an N terminal amino acid sequence given by SEQ ID No: 1. A short peptide was synthesized having SEQ ID No: 1. The synthetic peptide having the sequence SEQ ID No: 1 shows lethal toxin neutralizing activity similar to the natural **LTNF** from opossum or mongoose sera. The synthetic **LTNF** also has life saving utility.

ST opossum serum **lethal toxin neutralizing factor**

IT Poisoning (biological)  
RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
(botulism; opossum serum-derived **lethal toxin neutralizing factor** for universal treatment of snakebite or other envenomation and intoxication)

IT Peptides, biological studies  
RL: PRP (Properties); PUR (Purification or recovery); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(**lethal toxin neutralizing factors** (250-2,500 mol. wt.); opossum serum-derived **lethal toxin neutralizing factor** for universal treatment of snakebite or other envenomation and intoxication)

IT Allergies  
Biological warfare agents  
Chemical warfare agents

- Crotalus adamanteus  
 Crotalus atrox  
 Didelphidae  
 Didelphis virginiana  
 Elapidae  
 Hydrophidae  
 Inflammation  
 Naja naja atra  
 Naja naja kaouthia  
 Oxyuranus scutellatus  
 Protein sequences  
 Sepsis  
 Serum (blood)  
 Venoms  
 Vipera russelli  
 Viperidae  
 (opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT Ricins  
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
 (opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT Toxins  
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
 (plant or bacterial; opossum serum-derived **lethal**  
**toxin neutralizing factor** for universal  
 treatment of snakebite or other envenomation and intoxication)
- IT Snake  
 (poisonous; opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT Bee  
 Scorpion  
 (sting; opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT Plant (Embryophyta)  
 (toxin; opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT Bacteria (Eubacteria)  
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
 (toxin; opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT 9001-84-7, Phospholipase A2 9007-40-3, Crotoxin 12584-83-7, Cobratoxin  
 52019-39-3, Taipoxin  
 RL: ADV (Adverse effect, including toxicity); BIOL (Biological study)  
 (opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT 51-45-6, Histamine, biological studies  
 RL: ADV (Adverse effect, including toxicity); BSU (Biological study,  
 unclassified); BIOL (Biological study)  
 (opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)
- IT 184782-42-1 206755-63-7 206755-64-8  
 RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES  
 (Uses)  
 (opossum serum-derived **lethal toxin**  
**neutralizing factor** for universal treatment of  
 snakebite or other envenomation and intoxication)

L20 ANSWER 3 OF 3 HCPLUS COPYRIGHT 1999 ACS  
 AN 1996:713649 HCPLUS  
 DN 126:27960  
 TI Embodiments of natural and synthetic **lethal toxin neutralizing factors** and their utility as treatment for envenomation  
 IN **Lipps, Binie V.; Lipps, Frederick W.**  
 PA Lipps; Binie V., USA; Lipps; Frederick W.  
 SO U.S., 9 pp. Cont.-in-part of U.S. Ser. No. 53,387, abandoned.  
 CODEN: USXXAM

DT Patent  
 LA English  
 IC ICM A61K038-10  
 ICS A61K038-16; C07K007-08; C07K014-47  
 NCL 514014000  
 CC 4-5 (Toxicology)  
 FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 5576297	A	19961119	US 1994-310340	19940922
	US 5744449	A	19980428	US 1996-657163	19960603

PRAI US 1993-58387 19930510  
 US 1994-310340 19940922

AB Opossum whole serum exhibits a life saving property by **neutralizing** the **lethality** of venoms from all major families of poisonous snakes, and therefore an injection of Opossum serum can be used as a novel treatment for many types of envenomation. Preferably, the injectable treatment for envenomation should be a compn. obtained from the fraction of Opossum whole serum which contains the **lethal toxin neutralizing factor**, i.e. the so called natural **LTNF**, in purity. A method is given for the manuf. of a **lethal toxin neutralizing factor** from the serum of an opossum (*Didelphis virginiana*), by fractionating the opossum serum and isolating this select fraction from the plurality of fractions having an N terminal amino acid sequence given by SEQ ID No: 1. A short peptide was synthesized having SEQ ID No: 1. The synthetic peptide having sequence SEQ ID No: 1 shows **lethal toxin neutralizing** activity similar to the natural **LTNF** from opossum or mongoose sera. The synthetic **LTNF** also has life saving utility.

ST opossum **lethal toxin neutralizing factor** envenomation; **Didelphis lethal toxin neutralizing factor** envenomation

IT Coagulation **factors** (blood)  
 RL: PUR (Purification or recovery); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (**lethal toxin neutralizing factor**  
 ; natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for envenomation)

IT *Didelphis virginiana*  
 Serum (blood)  
 (**lethal toxin neutralizing factors** from opossum serum and utility as treatment for envenomation)

IT Antivenoms  
 Snake  
 (natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for envenomation)

IT 184782-42-1P  
 RL: PUR (Purification or recovery); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (natural and synthetic **lethal toxin neutralizing factors** and utility as treatment for

envenomation)

=> fil wpids  
FILE 'WPIDS' ENTERED AT 15:33:02 ON 29 DEC 1999  
COPYRIGHT (C) 1999 DERWENT INFORMATION LTD

FILE LAST UPDATED: 21 DEC 1999 <19991221/UP>  
>>>UPDATE WEEKS:  
MOST RECENT DERWENT WEEK 199954 <199954/DW>  
DERWENT WEEK FOR CHEMICAL CODING: 199954  
DERWENT WEEK FOR POLYMER INDEXING: 199954  
DERWENT WORLD PATENTS INDEX SUBSCRIBER FILE, COVERS 1963 TO DATE

>>> D COST AND SET NOTICE DO NOT REFLECT SUBSCRIBER DISCOUNTS -  
SEE HELP COST <<<

>>> IMPORTANT DERWENT ANNOUNCEMENT ABOUT CHANGES TO CPI  
SUBSCRIBER INDEXING - SEE NEWS <<<

>>> FOR UP-TO-DATE INFORMATION ABOUT ALL 'NEW CONTENT' CHANGES TO  
WPIDS, INCLUDING THE DERWENT CHEMISTRY RESOURCE (DCR),  
PLEASE VISIT <http://www.derwent.com/newcontent.html> <<<

=> d all abeq tot 130

L30 ANSWER 1 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD  
AN 1998-271108 [24] WPIDS  
CR 1997-011287 [01]  
DNC C1998-084507  
TI **Lethal Toxin Neutralising Factor**  
peptide from opossum - can neutralise venom(s) from all major families of  
poisonous snakes.  
DC B04  
IN **LIPPS, B V; LIPPS, F W**  
PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W  
CYC 1  
PI US 5744449 A 19980428 (199824)\* 11p A61K038-10  
ADT US 5744449 A CIP of US 1993-58387 19930510, CIP of US 1994-310340  
19940922, US 1996-657163 19960603  
FDT US 5744449 A CIP of US 5576297  
PRAI US 1996-657163 19960603; US 1993-58387 19930510; US 1994-310340  
19940922  
IC ICM A61K038-10  
ICS A61K038-16; C07K007-08; C07K014-47  
AB US 5744449 A UPAB: 19980617  
A new peptide (I) has a molecular weight of 250 to 2500 g/mole and  
comprises at least a 3 amino acid portion of:  
Leu-Lys-Ala-Met-Asp-Pro-Thr-Pro-Leu-Trp-Ile-Lys-Thr-Glu (A).  
Also claimed are:  
(1) a peptide as (I), but comprising at least a 5 amino acid portion  
of (A);  
(2) a method for treating a victim of envenomation from a poisonous  
snake comprising administering (I) to a victim by a route and in an amount  
which is effective to neutralise at least one effect of the envenomation;  
and  
(3) a method for treating a victim of envenomation from a poisonous  
snake family from the family of Elapidae, Viperidae or sea snake by  
administering a protein of about 68 kDa with an initial 15 amino acid  
sequence as (A) by a route and in an amount which is effective to  
neutralise at least one effect of the envenomation.  
USE - The peptide of (I) and the methods is a **Lethal**  
**Toxin Neutralising Factor** (LTNF) moiety from a  
68 kDa anti-haemorrhagic protein derived from an opossum. It is useful for  
the treatment of snake bites, sepsis, allergies caused by the environment

and treatment of bee or scorpion stings or toxicities caused by plant or bacterial toxins. (I) can also be used in histamine reaction treatment (claimed).

**ADVANTAGE** - The peptides can be used in envenomation treatment for a variety of snakes without prior identification of the guilty snake. Being short it can be synthetically prepared rather than the current production in horses, where some people can show hypersensitivity to horse proteins.

Dwg.0/3

FS CPI  
FA AB; DCN  
MC CPI: B04-C01C; B04-N04A; B14-A01; B14-G02A; B14-M01

L30 ANSWER 2 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD  
AN 1997-372086 [34] WPIDS  
DNC C1997-119850  
TI Topical treatment of herpes simplex virus infections - with combination of snake venom proteins.

DC B04  
IN LIPPS, B V; LIPPS, F W  
PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W  
CYC 1  
PI US 5648339 A 19970715 (199734)\* 8P A61K038-16  
ADT US 5648339 A US 1994-280157 19940725  
PRAI US 1994-280157 19940725  
IC ICM A61K038-16  
ICS A61K035-58; A61K038-17  
AB US 5648339 A UPAB: 19970820

Treatment of herpes simplex virus (HSV-1 or HSV-2) infections comprises topically applying to an affected area a composition comprising two phospholipase A2 proteins obtained from the snake venom of Naja naja kaouthia. The proteins have molecular weights of 14 and 15 kD and have the N-terminal amino acid sequence Asn-Leu-Tyr-Gln-Phe-Lys-Asn-Met-Ile-Gln-Cys-Thr-Val-Pro-Asn.

In an example, Chang's liver cell cultures were infected with HSV-1 or HSV-2 at 10<sup>-2</sup> to 10<sup>-6</sup> dilutions. After an absorption period, excess inoculum was removed and 1 ml of medium was added to each well. A 1:1 mixture of the two proteins was added at a concentration of 5 μg/ml. In all cases, no cytopathic effect was observed after 4 days.

**ADVANTAGE** - The composition inhibits the cytopathic effect of HSV-1 and HSV-2 at concentrations as low as 5 μg/ml without being cytotoxic.

Dwg.0/3

FS CPI  
FA AB; DCN  
MC CPI: B04-N02; B14-A02A3

L30 ANSWER 3 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD  
AN 1997-011287 [01] WPIDS  
CR 1998-271108 [24]  
DNC C1997-003058  
TI Treatment of victims of bee or scorpion stings or plant or bacterial toxins - by admin. of **lethal toxin-neutralising factor** or its N-terminal peptide.

DC B04 D16  
IN LIPPS, B V; LIPPS, F W  
PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W  
CYC 1  
PI US 5576297 A 19961119 (199701)\* 9P A61K038-10  
ADT US 5576297 A CIP of US 1993-58387 19930510, US 1994-310340 19940922  
PRAI US 1994-310340 19940922; US 1993-58387 19930510  
IC ICM A61K038-10  
ICS A61K038-16; C07K007-08; C07K014-47  
AB US 5576297 A UPAB: 19980617

Methods for treating victims of bee stings, scorpion stings, plant toxins or bacterial toxins comprise intravenous injection of (a) a **lethal-toxin-neutralising factor** (LTNF) obtained from an animal having resistance to envenomation, provided that the animal

is a Didelphis opossum in the case of bee and scorpion stings, or (b) a peptide of formula (I): Leu-Lys-Ala-Met-Asp-Pro-Thr-Pro-Pro-Leu-Trp-Ile-Lys-Thr-Glu (I). Also claimed is a method for treating a victim of a bee sting, comprising topical admin. of a soln. of (I).

ADVANTAGE - Use of natural or synthetic LTNF as a treatment for snake bites overcomes the problem of hypersensitivity occurring with horse-derived antivenom.

Dwg.0/3

FS CPI  
FA AB; DCN  
MC CPI: B04-C01C; B04-H01; B14-G02A; B14-M01; D05-H13

L30 ANSWER 4 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD  
AN 1996-476328 [47] WPIDS  
DNC C1996-148746

TI Treatment of cancers - using Atroporin or Kaotree or a combination, which are purified from the venom of snakes..

DC B04

IN LIPPS, B V; LIPPS, F W

PA (LIPP-I) LIPPS B V; (LIPP-I) LIPPS F W

CYC 1

PI US 5565431 A 19961015 (199647)\* 20p A61K038-17

ADT US 5565431 A CIP of US 1994-262678 19940620, US 1994-310341 19940922

PRAI US 1994-310341 19940922; US 1994-262678 19940620

IC ICM A61K038-17

ICS C07K014-46

AB US 5565431 A UPAB: 19961124

A claimed method for treating a patient having a cancer tumour comprising cancer cells comprises contacting the cancer cells with molecules of Atroporin and/or Kaotree at a dosage level sufficient to slow or regress the cancer tumour without significant destruction of normal cells which may also be contacted by Atroporin and/or Kaotree. The Atroporin and Kaotree each consist of a single purified fraction obtd. from snake venom, each fraction contg. a single peptide having either its first fifteen N-terminal amino acids as in (I) and a mol. wt. of about 35000 Daltons (for Atroporin) or having its first fifteen N-terminal amino acids as in (II) and a mol. wt. of about 6000 Daltons (for Kaotree).

Xaa-Xaa-Gly-Gly-Asp-Glu-Cys-Asn-Ile-Asn-Glu-His-Arg-Ser-Leu (I)

Met-Glu-Cys-Tyr-Arg-Met-Ser-Asn-Ile-Val-Thr-Cys-Gln-Pro-Trp (II)

USE - Atroporin and/or Kaotree show cytolytic activity against a wide range of tumour cells and so are useful as anticancer agents against e.g. cancers of the breast, colon, liver, lung, pharynx and ovary.

ADVANTAGE - The peptides do not cause the typical side effects caused by usual chemotherapy.

Dwg.0/7

FS CPI  
FA AB; DCN  
MC CPI: B04-N02B; B14-H01B

L30 ANSWER 5 OF 5 WPIDS COPYRIGHT 1999 DERWENT INFORMATION LTD

AN 1995-393068 [50] WPIDS

DNC C1995-169353

TI Use of beta-taipoxin as cell growth factor and mitogen - partic. to promote cell growth in serum free media and to improve healing of wounds.

DC B04 D16

IN LIPPS, B V

PA (LIPP-I) LIPPS B V

CYC 61

PI WO 9529987 A1 19951109 (199550)\* EN 20p C12N005-06

RW: AT BE CH DE DK ES FR GB GR IE IT KE LU MC MW NL OA PT SD SE SZ UG

W: AM AU BB BG BR BY CA CN CZ FI GE HU JP KG KP KR KZ LK LR LT LV MD

MG MN MX NO NZ PL RO RU SI SK TJ TT UA US UZ VN

AU 9526360 A 19951129 (199609) C12N005-06

EP 759069 A1 19970226 (199714) EN C12N005-06

R: BE CH DE FR GB IT LI SE

AU 691621 B 19980521 (199832) C12N005-06

ADT WO 9529987 A1 WO 1995-US5657 19950503; AU 9526360 A AU 1995-26360  
19950503; EP 759069 A1 EP 1995-921232 19950503, WO 1995-US5657 19950503;  
AU 691621 B AU 1995-26360 19950503

FDT AU 9526360 A Based on WO 9529987; EP 759069 A1 Based on WO 9529987; AU  
691621 B Previous Publ. AU 9526360, Based on WO 9529987

PRAI US 1994-237129 19940503

REP 06Jnl.Ref; US 4443546

IC ICM C12N005-06

ICS A61K038-18; C12N005-08

AB WO 9529987 A UPAB: 19951215

A culture medium contg. a peptide (I) having the N-terminal sequence (Ia)  
is new Asn-Leu-Val-Glu-Phe-Gly-Lys-Met-Ile-Glu-Cys-Ala-Ile-Arg-Asn  
(Ia).

USE - (I) is partic. beta-taipoxin isolated from snake venom. (I) is  
a cell growth factor and potent mitogen. It provides good cell growth in  
absence of serum (e.g. for growth of skin monolayers) and provides rapid,  
scar-free healing of cuts, burns and abrasions (when applied topically).

ADVANTAGE - Cells can now be grown effectively in serum-free medium  
so purification of cellular products is facilitated. At 0.1 mug/ml (Ia)  
has an effect equiv. to 10% serum for a wide range of cells. (Ia) is water  
soluble, stable at 4deg.C or room temp. and not toxic to cells at concns.  
of use.

Dwg.0/3

FS CPI

FA AB; DCN

MC CPI: B04-C01C; B04-F02; B14-N17B; D05-H08; D05-H17A2

=> d his 131-

FILE 'WPIDS' ENTERED AT 15:33:02 ON 29 DEC 1999

L31 2 S LEU LYS ALA MET ASP PRO THR PRO PRO LEU TRP ILE LYS THR GLU  
L32 0 S L31 NOT L30

Wed Dec 29 14:17:34 1999

baskar-09-300612.pep

Page 1

; Entered [jdejava1 29-Dec-99 14:04]  
09-300612  
lkamoptppiwktel

\*\*\*\*\*

KW histamine reaction treatment.

OS Didelphis virginiana.

PN

US5744441-A.

PD 28-APR-1998.

PR 03-JUN-1996; 657163.

PR 03-JUN-1996; US-657163.

10-MAY-1993; US-058387.

PR 22-SEP-1994; US-310340.

PA (LIPPP) LIPPS B V.

PA (LIPPP) LIPPS F W.

PI LIPPS BV, LIPPS FW;

DR WPI; 98-271108/24.

PT Lethal Toxin Neutralising Factor peptide from opossum - can

neutralise venom(s) from all major families of poisonous snakes

PS Claim 7; Column II; LIPP; English.

CC This sequence represents the peptide of the invention. It is a Lethal

Toxin Neutralising Factor (LTNF) moiety from a 68 kDa anti-haemorrhagic

protein derived from an opossum.

CC treating a victim of envenomation from a poisonous snake, preferably a

CC poisonous snake from the family of Elapidae, Viperidae or sea snake.

CC It is useful for the treatment of snake bites, sepsis, allergies caused

CC by the environment and treatment of bee or scorpion stings or toxicities

CC caused by plant or bacterial toxins. The peptide can also be used in

CC histamine reaction treatment. The peptide can be used in envenomation

CC treatment for a variety of snakes without prior identification of the

CC snake. Being short it can be synthetically prepared rather than the

CC current production in horses, where some people can show hypersensitivity

CC to horse proteins.

SQ Sequence 10 AA;

Query Match 100.0%; Score 73; DB 30; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.13e-01; Matches 10; Mismatches 0; Indels 0; Gaps 0;  
 Job time : 91 secs.  
 Db 1 lkamdpptpl 10  
 Qy 1 lkamdpptpl 10  
 |||||||||

Statistics:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

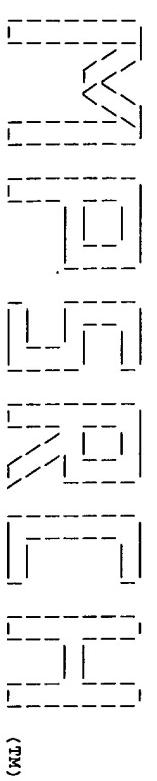
Result No.	Score	Query Match Length	DB ID	Description	Pred. No.
1	73	100.0	10 30	W53843 N-terminus of opossum	3.13e-01

Note: Post-processor removed 999 summaries from list due to search parameters chosen.

#### ALIGNMENTS

RESULT	1	ID	W53843	standard; peptide; 10 AA.
AC				
DT	08-JUL-1998 (first entry)			
DE	N-terminus of opossum LTNF.			
KW	LTNF: lethal toxin neutralising factor; opossum; envenomation; therapy; anti-haemorrhagic protein; Elapidae; Viperidae; sea snake; snake bite; sepsis; allergy; bee sting; scorpion sting; plant toxin; bacterial toxin;			

\*\*\*\*\*  
\*\*\*\*\*  
\*\*\*\*\*  
\*\*\*\*\*  
\*\*\*\*\*



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:51:05 1999; MasPar time 1.65 Seconds  
72.120 Million cell updates/sec  
Tabular output not generated.

Title: Description: >09-300612  
(1-10) from baskar-09-300612.pep  
Perfect Score: 73  
Sequence: 1 lkamdpptppl 10

Scoring table: PAM 150

Searched: 122461 seqs, 11912985 residues

Post-processing: Minimum Match 100%  
Listing first 1000 summaries  
Maximum DB seq length 10

Database: a-issued

1:5A\_COMB 2:5B\_COMB 3:PCT9\_COMB 4:backfiles1

Statistics: Mean 15.231; Variance 49.318; scale 0.309

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description	Pred. No.
1	73	100.0	10	1 US-08-657- Sequence 2, Application 1.53e-01	

Note: Post-processor removed 999 summaries from list due to search parameters chosen.

#### ALIGNMENTS

RESULT 1 ID US-08-657-163A-2 STANDARD: PRT; 10 AA.

ID	AC	XX	DT	DE
XX	xxxxxx			

Sequence 2, Application US/08657163A  
Sequence 2, Application US/08657163A  
Patent No. 574449

GENERAL INFORMATION:  
APPLICANT: BINIE V. LIPPS AND FREDERICK W. LIPPS  
TITLE OF INVENTION: EMBODIMENTS OF NATURAL AND  
TITLE OF INVENTION: SYNTHETIC LIPINS AND THEIR

TITLE OF INVENTION: UTILITY AS TREATMENT FOR ENVENOMATION  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BINIE V. LIPPS  
STREET: 4505 MIMOSA DR.  
CITY: BELLAIRE  
STATE: TEXAS  
COUNTRY: USA  
ZIP: 77401

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" FLOPPY DISK, 1.44 MB  
COMPUTER: IBM COMPATIBLE  
OPERATING SYSTEM: MS-DOS 5.0/WINDOWS 3.1  
SOFTWARE: MS WORD 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/657,163A  
FILING DATE:  
CLASSIFICATION: 514  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/310,340  
FILING DATE: 22 SEPTEMBER 1994  
CLASSIFICATION: 514  
APPLICATION NUMBER: 08/058,387  
FILING DATE: 10 MAY 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: JOHN R. CASPERSON  
REGISTRATION NUMBER: 28,198  
REFERENCE/DOCKET NUMBER: FWL-PAT-US-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-482-2961  
TELEFAX: 713-663-7290  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE IN SEQ ID NO: 2  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N  
ORIGINAL SOURCE: SYNTHETIC  
SEQUENCE 10 AA; 1082 MW; 684 CN;  
SQ

Query Match 100.0%; Score 73; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.53e-01; 0; Mismatches  
Matches 10; Conservative 0; Indels 0; Gaps 0;  
Db 1 lkamdpptppl 10  
Qy 1 lkamdpptppl 10

Search completed: Wed Dec 29 15:51:50 1999  
Job time : 45 secs.

\* \* \* \* \*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

**MPSrch\_PP** protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:52:08 1999; MasPar time 8.77 Seconds

Tabular output not generated.

卷之三

**Title:** 209-300612  
**Description:** (1-10) from baskar-09-300612.pep  
**Default Name:** 73

Sequence: 1 lkamdpptpl 10  
Perfect score: ✓

**Scoring table:** PAM 150  
GAP 15

תְּמִימָנֶה וְעַמְּדָה וְעַמְּדָה וְעַמְּדָה

Sealine, 34-422 seqs., 1133421 residues

**post processing:** minimum match 100%  
Listing first 1000 summaries  
Maximum document length 10

[Pending](#) [U60](#) [U7](#) [U80](#) [U81](#)

10:00 11:00 12:00 13:00  
18:U93 19:NEWP 20:NEWU6 21:N

Statistics: Mean 17.858; Variance 49.093;

Pred. No. is the number of results predicted greater than or equal to the score

score greater than or equal to the score and is derived by analysis of the totals

SUMMARIES

Result  
No  
Score  
Query  
Watch  
Search  
Rank  
Recommender  
Name  
Version

No.	Score	Match	Length	DB	ID	Description	Freq.	No.
No matches found.								

Search completed: Wed Dec 29 15:53:52 1999

Job time : 104 secs.

Seal City Complicated

## SUMMARIES

\* \* \* \* \*

A vertical column of five dashed-line mazes, each with a single solution path. The first maze is a small hexagon. The second is a square with a central cross. The third is a larger square with a central T-shape. The fourth is a rectangle with a diagonal line from top-left to bottom-right. The fifth is a rectangle with a vertical line from top to bottom.

卷之三

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

HISTORICAL PERSPECTIVE ON THE DEVELOPMENT OF THE POLYMER INDUSTRY

108.566 Million cell updates/sec

卷之三

Perfect Score: 73 Description: 100% basketball - 0% - 000012.ppt

Gap 15

Searched: 122810 seqs, 40068593 residues

listing first 1000 summaries

1:pir1 2:pir2 3:pir3 4:pir4

STATISTICS: Mean 23.01; Variance 29.210; Scale 0./88

the chance to have a score greater than or equal to the score of the results printed.

כוננותם

No matches found.

Search completed: Wed Dec 29 15:46:16 1999

LAWRENCE BROWN

A vertical column of five dashed-line mazes, each with a star at the top right corner. The mazes increase in complexity from left to right.

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.

**MPSrch\_PP** protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:48:23 1999; MasPar time 2.76 Seconds  
102.460 Million cell updates/sec

三三三三三三三三

Description: (1-10) from baskar-09-300612.pep  
Perfect Score: 73

Scoring table: PAM 150

post-processing:  
minimum batch 100%  
listing first 1000 summaries  
Maximum DB seq length 10

Database: swissprot37  
1:swissprot

Statistics: Mean 23.724; Variance 26.440; scale 0.897

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total sample distribution.

## SUMMARIES

## Result Query

NO.	Score	Match Length	DB	IU	Description	Pred.	No.
-----	-------	--------------	----	----	-------------	-------	-----

No matches found.

Search completed: Wed Dec 29 15:49:14 1999  
Job time : 51 secs.

\* \* \* \* \*

(四三)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

KODAK SAFETY FILM FOR STEREOVIEWERS

Run On: Wed Dec 30 15:49:32 1999. Macstar time 5 28 seconds

Tabular output not generated.

Title: >09-300612

Description: (1-10) from baskar-09-300612.pep  
Perfect Score: 73

Sequence: 1 lkamdpptpl 10

Gap 15

Searched: 179066 seqs, 54579741 residues

**Post-processing:** Minimum Match 100% listing float 1000 summarizing

Maximum DB seq length 10

Databasen sptrembly inarchea 3:en bacteria 3:

1:sp\_archea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
 5:sp\_invertebrate 6:sp\_mammal 7:sp\_micr 8:sp\_organelle  
 9:sp\_phage 10:sp\_plant 11:sn\_rordan 12:sn\_unclassified

13:SP-Ventilatorable 14:SP-Ventilator

Mean 22.1740, Variance 21.0151, Scale 0.0040

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the to

100

No.	Score	Match	Length	DB	ID	Description	Pred.	No.
-----	-------	-------	--------	----	----	-------------	-------	-----

No matches found

Search completed: Wed Dec 29 15:50:47 1999  
Job time : 75.000 secs



Page  
1

A vertical stack of five dashed-line mazes. Each maze consists of a rectangular frame with various internal paths and dead ends. The shapes to find are: a zigzag at the top, a square in the middle, a rectangle below it, a triangle further down, and a diamond at the bottom.

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

**Mpsrch\_pp** protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed Dec 29 15:50:54 1999; MasPar time 1.54 Seconds  
38.739 Million cell updates/sec

Title: >09-300612  
Description: (1-5) from baskar-09-300612.pep  
Perfect Score: 36

Sequence: 1 lkamd 5  
Scoring table: PAM 150  
Gap 15

Searched: 122461 seqs, 11912985 residues  
Post-processing: Minimum Match 100%  
Listing first 1000 summaries

Database: Maximum DB seq length 5  
a-issued 1:5A..COMB 2:5B..COMB 3:PCT9-

Statistics: Mean 12.206; Variance 31.991  
Pred. No. is the number of results predicted correctly according to the decision rule.

Score greater than or equal to the total score and is derived by analysis of the total

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	36	100.0	5	1	US-08-657 - Sequence 3, Application	1.73e+027

Note: Post-processor removed 822 summaries from list due to search parameters chosen.

RESULT 1 NOT AVAILABLE.  
ATTACHMENTS

RESULT 1  
ID US-08-657-163A-3  
XX STANDARD; PRT; 5 AA.

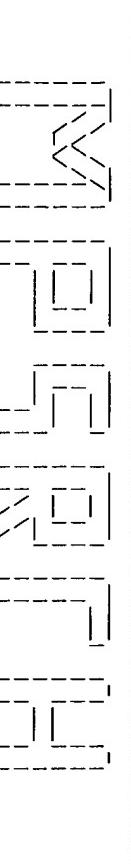
Sequence 3, Application US/08657163A

Sequence 3: Application US/08657163A

Patent No. 5744449

GENERAL INFORMATION





\*\*\*\*\*  
\*\*\*\*\*  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

MPSEARCH\_PP protein - Protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:46:58 1999; Maspar time 2.93 Seconds  
68.364 Million cell updates/sec

Tabular output not generated.

Title: >09-300612

Description: (1-5) from baskar-09-300612.pep

Perfect Score: 36

Sequence: 1 lkand 5

Scoring table: PAM 150

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 100%  
Listing first 1000 summaries  
Maximum DB seq length 5

Database: Pir60

1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 18.837; Variance 18.467; scale 1.020

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

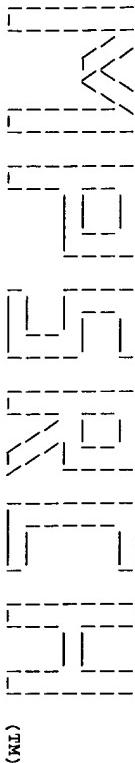
#### SUMMARIES

Result	Query	Score	Match Length	DB ID	Description	Pred. No.
--------	-------	-------	--------------	-------	-------------	-----------

No matches found.

Search completed: Wed Dec 29 15:47:54 1999  
Job time : 56 secs.





Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

**MPStch\_pp** protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 29 15:49:17 1999; MasPar time 4.11 Seconds

Tabular output not generated.

Title: >09-300612

Description: (1-5) from baskar-09-300612.pep

Perfect Score: 36

Sequence: 1 lkmd 5

Scoring table: PAM 150

Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 100%  
 Listing first 1000 summaries

Maximum DB seq length 5

Database:

sptrembl9  
 1:sp\_archea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
 5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
 9:sp\_phage 10:sp\_plant 11:sp\_rat 12:sp\_unclassified  
 13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 18.536; variance 16.495; scale 1.124

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description	Pred. No.
------------	-------	--------------------	-------	-------------	-----------

No matches found.

Search completed: Wed Dec 29 15:50:37 1999  
 Job time : 80 secs.

\*\*\*\*\*

Release 3.1A, John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1995-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

**protein - protein database search, using Smith-Waterman algorithm**

run on: Wed Dec 29 14:06:35 1999; Maspar time 2.57 Seconds  
164.988 Million cell updates/sec

ribular output not generated.

title:  
scription: >-9-300612  
(1-15) from baskar-09-300612.pep

perfect Score: 115  
quence: 1lkandptpplwkte 15

scoring table: PAM 150  
Gap 15

searched: 77977 seqs, 28268293 residues

post-processing: Minimum Match 0%  
Listing first 45 summaries

atabase: swiss-prot37  
1:swissprot

**statistics:**  
Mean 27.347; Variance 37.123; scale 0.737

**Pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

result	1	standard:	prt:	1213 aa.
ID	T2D2_DROME			
AC	Q24225;			
DT	01-NOV-1997	(REL. 35, CREATED)		
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)		
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	TRANSCRIPTION INITIATION FACTOR TFIID 150 KD SUBUNIT (TAFFI150).			
DE	(TAFFI150).			
GN	TAFFI150.			
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).			
OC	EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA; PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA; DROSOPHILIDAE; DROSOPHILA.			
OC	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE-EMBRYO;			
RA	VERBIZER C.P., YOKOMORI K., CHEN J.-L., TUJIAN R.;			
RT	*Drosophila TAFFI150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA.";			
RL	SCIENCE 264:933-941(1994).			
CC	-1- FUNCTION: TAFS ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA POLYMERASE TRANSCRIPTION. TAFFI-150 IS AN ESSENTIAL SUBUNIT WHICH INTERACTS DIRECTLY WITH TBP AND TAFFI-250 AND Binds TO CORE PROMOTOR DNA.			
CC	-1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A NUMBER OF TBP-ASSOCIATED FACTORS (TAFS).			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-1- SIMILARITY: TO YEAST TAFFI-150. (TSM1).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
DR	EMBL: X9243; G541665; -			
DR	TRANSFAC: T0120; -			
KW	TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.			
FT	DOMAIN 845 1213 BINDS TO TBP AND TAFFI-250.			
FT	DOMAIN 1138 1183 HIGHLY CHARGED.			
SQ	SEQUENCE 1213 AA; 138533 MW; 2A07DC14 CRC32;			

\*\*\*\*\*

Query Match Score 72; DB 1; Length 1213;  
 Best Local Similarity 53.3%; Pred. No. 6.17e-03;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 620 LSAMDSPVWIRD 634 QY 1 lkandptppplw1 12

RESULT 2 ID THTR\_AZOVI STANDARD; PRT; 271 AA.

ID THTR\_AZOVI STANDARD; PRT; 271 AA.

AC P52197; DT 01-OCT-1996 (REL. 34, CREATED)

AC P52197; DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

OC AZOTOBACTER. DT 15-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

RN [1] DE THIOSULFATE SULFURTRANSFERASE (EC 2.8.1.1) (RHODANESE-LIKE PROTEIN).

DE RHDA.

OS AZOTOBACTER VINELANDII. RN

OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; AZOTOBACTERACEAE; AZOTOBACTER.

OC SCHISTOCERCA. RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.

RN STRAIN=OP / UW136; RX MEDLINE: 9614904.

RA COTANGHI R., PAGANI S., KENNEDY C., DRUMMOND M.; RT Cloning, sequence analysis and overexpression of the rhodanese gene of Azotobacter vinelandii.

RL EUR. J. BIOCHEM. 236:240-248(1995).

RN [2] CC -I - CATALYTIC ACTIVITY: THIOSULFATE + CYANIDE = SULFITE + THIOCYANATE.

RX MEDLINE: 8822351.

RA ZINN K., MCALLISTER L., GOODMAN C.; CC -I - SIMILARITY: BELONGS TO THE RHODANESE FAMILY.

RT Sequence analysis and neuronal expression of fasciclin I in grasshopper and Drosophila.";

RL CELL 53:577-587(1988).

RN [2] CC -----

RX MEDLINE: 88226943.

RA SNOW P.M., ZINN K., HARRELSON A.L., MCALLISTER L., SCHILLING J., CC -----

RA BASTIANI M.J., MAKK G., GOODMAN C.S.; CC -----

RT Characterization and cloning of fasciclin I and fasciclin II CC -----

RT glycoproteins in the grasshopper.;" CC -----

RL PROC NATL ACAD SCI U.S.A. 85:5251-5255(1988). CC -----

CC -I - FUNCTION: NEURAL CELL ADHESION MOLECULE.

CC -I - SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.

CC -I - TISSUE SPECIFICITY: EXPRESSED ON DIFFERENT SUBSETS OF AXON BUNDLES (FASCICLES) IN INSECT EMBRYOS.

CC -----

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; PS00683; RHODANESE\_1; 1. CC -----

DR PROSITE; PS00683; RHODANESE\_2; 1. CC -----

DR PFAM; PF00581; RHODANESE\_1. CC -----

DR PROSITE; PS00683; RHODANESE\_2. CC -----

DR TRANSFERASE. FT ACT SITE 230 BY SIMILARITY.

SQ SEQUENCE 271 AA; 29829 MW; 6C39E3B CRC32;

Query Match Score 62; DB 1; Length 271;  
 Best Local Similarity 53.8%; Pred. No. 6.08e-01;  
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 198 AMDPSRALRIRD 210 QY 3 amdpptppplw1 15

RESULT 4 ID YC4\_TYLCA STANDARD; PRT; 102 AA.

ID YC4\_TYLCA STANDARD; PRT; 102 AA.

AC P36283; DT 01-JUN-1994 (REL. 29, CREATED)

AC P36283; DT 01-OCT-1994 (REL. 29, LAST SEQUENCE UPDATE)

DE HYPOTHETICAL 11.4 KD PROTEIN (C4 PROTEIN).

GN C4.

OS TOMATO YELLOW LEAF CURL VIRUS (STRAIN AUSTRALIA) (TYLCV).

OC VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; BEGOMOVIRUS.

RN [1] RP SEQUENCE FROM N.A.

RX MEDLINE: 93139778.

RA DRY I.B., RIGDEN J.E., KRAKE L.R., MULLINEAUX P.M., REZAIAN M.A.; RT "Nucleotide sequence and genome organization of tomato leaf curl virus.", J. GEN. VIROL. 74:147-151(1993).

RA LSELDENPPLWI 115 CC -----

Query Match Score 64; DB 1; Length 662;  
 Best Local Similarity 58.3%; Pred. No. 2.51e-01;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db • 104 LSELDENPPLWI 115 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: S53251; E96189; -.

CC PIR: JO1890; JO1890.

DR HYPOTHETICAL PROTEIN.

SQ SEQUENCE 102 AA; 11410 MW; 32D88BF CRC32;

Query Match 50.9%; Score 59; DB 1; Length 102;  
Best Local Similarity 40.0%; Pred. No. 2.22e+00;  
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 44 LRAROTSSSPWWRRE 58  
QY 1 ikandptppplwke 15

RESULT 5

ID DIV\_ECOLI STANDARD; PRT; 331 AA.

AC P15206; P77706; -.

DT 01-APR-1990 (REL. 14, CREATED)

DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE DIV PROTEIN.

GN DIV.

OS ESCHERICHIA COLI.

OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;

RC ESCHERICHIA.  
[1]

RN SEQUENCE FROM N.A.

RP STRAIN=K12 / MG1655;

RX MEDLINE; 97426617.

RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
RILEY M., COLLADO-VIDES J., RODE C.K., MAYHEW G.F.,  
GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
RA MAU B., SHAO Y.,

RT "The complete genome sequence of Escherichia coli K-12.";  
RL SCIENCE; 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,  
RA IKEMOTO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K., KASAI H.,  
RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,  
RA MASUDA S., MIKI T., MIZORUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,  
RA NASHIMOTO H., NISHIO Y., OSHIMA T., SATO N., SAMPEI G., SERI Y.,  
RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.,

RA SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / W3110;

RA PEASE A.J., SCHROENLEIN P.V., WINKLER M.E.,

RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.

RN [4]

RP SEQUENCE OF 1-39 FROM N.A.

RC STRAIN=K12;

RA MEDLINE: 90036695.

RA SCHROENLEIN P.V., ROA B.B., WINKLER M.E.;

RT "Divergent transcription of pdxB and homologs between the pdxB and  
serA gene products in Escherichia coli K-12.";

RL J. BACTERIOL. 171:6084-6092(1989).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ARCH: VIROL. 124:371-377(1992).

CC !- SIMILARITY: BELONGS TO THE ARENAVIRUSES NUCLEOCAPSID PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENTITIES REQUIRING A LICENSE AGREEMENT (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; M29962; E17525; -.

CC DR EMBL; M15541; -; NOT\_ANNOTATED\_CDS.

CC DR ECODINE; EG10229; DIV.

CC SQ SEQUENCE 331 AA; 36668 MW; D81EC9A4 CRC32;

Query Match 50.9%; Score 59; DB 1; Length 331;  
Best Local Similarity 50.0%; Pred. No. 2.22e+00;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 306 LQNMMATPAANI 317  
QY 1 ikandptppplwke 12

RESULT 6

ID NCAP\_MACHU STANDARD; PRT; 564 AA.

AC P26578;

DT 01-AUG-1992 (REL. 23, CREATED)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE NUCLEOCAPSID PROTEIN (NUCLEOPROTEIN).

GN MACHUPO VIRUS.

OS MACHUPO VIRUS; SSRNA NEGATIVE-STRAND VIRUSES; ARENAVIRIDAE; ARENAVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RA288-77;

RX MEDLINE; 92296904.

RA GRIFFITHS C., WILSON S.M., CLEGG J.C.S.;

RT "Sequence of the nucleocapsid protein gene of Machupo virus: close relationship with another South American pathogenic arenavirus, Junin.;"

RL ARCH: VIROL. 124:371-377(1992).

CC !- SIMILARITY: BELONGS TO THE ARENAVIRUSES NUCLEOCAPSID PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENTITIES REQUIRING A LICENSE AGREEMENT (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; X62616; G60522; -.

DR PIR: S18042; VHXPWY.

DR PFAM: PF00843; Arena\_nucleocap; 1.

DR KW NUCLEOCAPSID.

SQ SEQUENCE 564 AA; 63300 MW; 4BE0C9FA CRC32;

Query Match 50.9%; Score 59; DB 1; Length 564;  
Best Local Similarity 46.7%; Pred. No. 2.22e+00;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 368 VKKIDPPTWNLWIE 382  
QY 1 ikandptppplwke 15

RESULT 7

ID LAF4\_HUMAN STANDARD; PRT; 1227 AA.

AC P51026;

DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE LAF-4 PROTEIN (LYMPHOID NUCLEAR PROTEIN).

GN LAF4.

OS HOMO SAPIENTIS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 96141096.

RA MA C., STAUDT L.M.  
 RT \*LAF-4 encodes a lymphoid nuclear protein with transactivation  
 potential that is homologous to AF-4, the gene fused to ML in  
 t(4;11) leukemias.  
 RL BLOOD 87:734-745(1996).  
 CC -I- FUNCTION: PUTATIVE TRANSCRIPTION ACTIVATOR THAT MAY FUNCTION IN  
 LYMPHOID DEVELOPMENT AND ONCOGENESIS. BINDS, IN VITRO, TO  
 DOUBLE-STRANDED DNA.  
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -I- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LYMPHOID TISSUES.  
 CC -I- HIGHEST LEVELS BEING FOUND IN THE THYMUS.  
 CC -I- SIMILARITY: TO AF4 AND OX19.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to license@isb-sib.ch).

CC DR EMBL; U34360; G114493; -;  
 KW NUCLEAR PROTEIN; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.  
 CC FT DOMAIN 413 419 POLY-SER.  
 FT DOMAIN 422 432 POLY-SER.  
 FT DOMAIN 440 445 POLY-SER.  
 FT DOMAIN 670 679 POLY-SER.  
 SQ SEQUENCE 1227 AA; 133734 MW; 5F8C8BBB CRC32;

Query Match 50.9%; Score 59; DB 1; Length 1227;  
 Best Local Similarity 40.0%; Pred. No. 2.22e+00;  
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 757 LKDSBIRSLWKKD 771  
 Qy ||| :||| :  
 1 lkamdpptppiwkite 15

RESULT 8  
 ID Y4VJ RHISN STANDARD; PRT: 351 AA.  
 AC Q53218;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 39.2 KD PROTEIN Y4VJ.  
 GN Y4VJ.  
 OS RHIZOBIUM SP. (STRAIN NGR34).  
 OC PLASMID SYM PGR3234A.  
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
 OC RHIZOBIAEAE; RHIZOBIUM.  
 RN [1] SEQUENCE FROM N.A.  
 RX MEDLINE; 97305356.  
 RA FREIBERG C.A., FELLY R., BAIROCH A., BROUGHTON W.J., ROSENTHAL A.,  
 RP PERRET X.;  
 RA "Molecular basis of symbiosis between Rhizobium and legumes.",  
 RL NATURE 387:394-401(1997).  
 RN [2] SEQUENCE OF 1-279 FROM N.A.  
 RX MEDLINE; 96389014.  
 RA FREIBERG C., PERRET X., BROUGHTON W.J., ROSENTHAL A.,  
 RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.  
 NGR3234 using dye terminators and a thermostable 'sequenase'; a  
 beginning";  
 RL GENOME RES. 6:590-600(1996).  
 CC -I- SIMILARITY: TO ALKANAL MONOXYGENASE ALPHA AND BETA CHAINS (EC  
 1.14.14.3) (BACTERIAL LUCIFERASE), BUT DISTANTLY RELATED.  
 CC -I- SIMILARITY: TO Y4VJ.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to license@isb-sib.ch).

RESULT 9  
 ID RPL-LAHSV9 STANDARD; PRT: 1305 AA.  
 AC 070655;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (VP1).  
 GN S1.  
 OS AFRICAN HORSE SICKNESS VIRUS 9 (AHSV-9) (AFRICAN HORSE SICKNESS VIRUS  
 OS (SEROTYPE 9)).  
 OC VIRUSES; DSRNA VIRUSES; REOVIRIDAE; ORBIVIRUS.  
 RN [1] SEQUENCE FROM N.A.  
 RX MEDLINE; 98202706.  
 RA VREEDA F.T., HUTSMANS H.;  
 RP "Sequence analysis of the RNA polymerase gene of African horse  
 RT sickness virus.";  
 RL ARCH VIROL. 143:113-419(1998).  
 CC -I- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE  
 CC + RNA(N).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to license@isb-sib.ch).

CC DR EMBL; U94887; G2961464; -;  
 KW TRANSFERASE; NUCLEOTIDYL-TRANSFERASE; RNA-DIRECTED RNA POLYMERASE.  
 SQ SEQUENCE 1305 AA; 150293 MW; 251BD435 CRC32;

Query Match 50.0%; Score 58; DB 1; Length 1305;  
 Best Local Similarity 50.0%; Pred. No. 3.38e+00;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 960 VKLIDQSPPLNET 973  
 Qy 1 lkamdpptppiwkite 14

RESULT 10  
 ID CO4 HUMAN STANDARD; PRT: 1741 AA.  
 AC P01058;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 13-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE COMPLEMENT C4 PRECURSOR [CONTAINS: C4A ANAPHYLATOXIN].  
 GN CAA AND C4B.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; MAMMALIA; EUTHERIA;  
 PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]

RP SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.  
 MEDLINE: 85156269.  
 RA BELT K.T., YU C.Y., CARROLL M.C., PORTER R.R.;  
 RT "POLYMORPHISM of human complement component C4.",;  
 RL IMMUNOGENETICS 21:173-180(1985).  
 RN [2]  
 RP SEQUENCE OF 20-1741 FROM N.A.  
 RC TISSUE-LIVER.  
 RX MEDLINE; 84156544.  
 RA BELT K.T., CARROLL M.C., PORTER R.R.;  
 RT "The structural basis of the multiple forms of human complement  
 component C4";  
 RT "complete primary structure of human C4a anaphylatoxin.",;  
 RL J. BIOL. CHEM. 256:8685-8692(1981).  
 RN [4]  
 RP SEQUENCE OF 957-1444.  
 RX MEDLINE; 82102029.  
 RA CAMPBELL R.D., GAGNON J., PORTER R.R.;  
 RT "Amino acid sequence around the thiol and reactive acyl groups of  
 human complement component C4.",;  
 RL BIOCHEM. J. 199:359-370(1981).  
 RN [5]  
 RP SEQUENCE OF 990-1017.  
 MEDLINE; 82150875.  
 RX HARRISON R.A., THOMAS M.L., TACK B.F.;  
 RA SARGENT C.A., ANDERSON M.J., HSIEH S.L., KENDALL E.,  
 RT component of human complement.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 78:7388-7392(1981).  
 RN [6]  
 RP SEQUENCE OF 1-21 FROM N.A.  
 RX MEDLINE; 94282044.  
 RA GOMEZ-ESCORBAR N., CAMPBELL R.D.;  
 RT "Characterisation of the novel gene G11 lying adjacent to the  
 complement C4A gene in the human major histocompatibility complex.",;  
 RL HUM. MOL. GENET. 3:481-488(1994).  
 RN [7]  
 RP STRUCTURAL BASIS OF POLYMORPHISM.  
 RX MEDLINE; 87050272.  
 RA YU C.Y., BELT K.T., GILES C.M., CAMPBELL R.D., PORTER R.R.;  
 RT "Structural basis of the hemolytically inactive C4a6 allotype of  
 C4A and C4B: gene size, reactivity and antigenicity.",;  
 RL EMBO J. 5:2873-2881(1986).  
 RN [8]  
 RP VARIANT C4A6 ALLOTYPE.  
 RX MEDLINE; 92242905.  
 RA ANDERSON M.J., MILLER C.M., COTTON G.H., CAMPBELL R.D.;  
 RT "The coding sequence of the hemolytically inactive C4a6 allotype of  
 human complement component C4 reveals that a single arginine to  
 tryptophan substitution at beta-chain residue 458 is the likely cause  
 of the defect";  
 RL J. IMMUNOL. 148:2795-2802(1992).  
 CC -1- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY  
 ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A  
 ANAPHYLATOXIN.  
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,  
 C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
 INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
 PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
 BASOPHILIC LEUKOCYTES.  
 CC -1- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR  
 AND, PRIOR TO SECRETION, IS ENZYMICALLY CLEAVED TO FORM A TRIMER  
 OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).  
 CC -1- POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT  
 LEAST TWO LOCI C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF  
 C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.  
 CC -1- POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE

CC THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.  
 CC -1- DISEASE: THE C4A6 ALLOTYPE IS TOTALLY DEFICIENT IN HEMOLYTIC  
 ACTIVITY.  
 CC -1- C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO GROUP OF PEPTIDE  
 ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY WITH THE HYDROXYL  
 GROUP OF CARBOHYDRATE ANTIGENS.  
 CC -1- C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III PROTEIN.  
 CC -1- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; M14823; G553211; -.  
 DR EMBL; K02403; G443671; ALT\_SEQ.  
 CC DR EMBL; M14824; G553210; -.  
 DR EMBL; X7791; G453411; -.  
 DR PIR; A01262; C4HU.  
 DR PIR; A17265; A17265.  
 DR PIR; A29177; A29177.  
 DR PIR; B20807; B20807.  
 DR SWISS-2DPAGE; P01028; HUMAN.  
 DR MIM; 12070; -.  
 DR MIM; 120810; -.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 DR PRIM; PF00207; AZM; 1.  
 DR HSSP; P01031; ICKS.  
 KW COMPLEMENT PATHWAY; PLASMA; GLYCOPROTEIN; MHC III; SIGNAL;  
 KW INFLAMMATORY RESPONSE; POLYMORPHISM; DISEASE MUTATION;  
 KW BLOOD GROUP ANTIGEN.  
 FT SIGNAL 1 19  
 FT CHAIN 20 675  
 FT PROPEP 676 679  
 FT CHAIN 680 1443  
 FT PROPEP 1444 1450  
 FT CHAIN 1451 1741  
 FT PEPTIDE 756 680  
 FT DOMAIN 702 736  
 FT DISULFID 702 728  
 FT DISULFID 703 735  
 FT DISULFID 716 736  
 FT THIOLEST 1010 1013  
 FT CARBOHYD 226 226  
 FT CARBOHYD 862 862  
 FT CARBOHYD 1328 1328  
 FT CARBOHYD 1391 1391  
 FT VARIANT 477 477  
 FT VARIANT 726 726  
 FT POTENTIAL.  
 FT VARIANT 1073 1073  
 FT C4A1, C4B1 AND C4B3).  
 FT PCVLD -> LSPVIIH (IN C4B).  
 FT VARIANT 1120 1125  
 FT N -> S (IN C4A1, C4B1, C4B3 AND C4B5).  
 FT VARIANT 1176 1176  
 FT S -> T (IN C4A6, C4A3, C4A1, AND C4B).  
 FT VARIANT 1201 1201  
 FT VARIANT 1207 1207  
 FT V -> A (IN C4A1, C4B1, C4B2 AND C4B3).  
 FT VARIANT 1210 1210  
 FT L -> R (IN C4A1, C4B1, C4B2 AND C4B3).  
 FT VARIANT 1286 1286  
 FT S -> A (IN C4A6, C4A1, C4A3 AND C4B).  
 FT CONFLICT 727 727  
 FT D -> N (IN REF.).  
 FT CONFLICT 1013 1013  
 FT Q -> E (IN REF. 4 AND 5).  
 SQ SEQUENCE 1741 AA; 192335 MW; 402A8805 CRC22;  
 Query Match 50.0%; Score 53; DB 1; Length 1741;  
 Best Local Similarity 50.0%; Pred. No. 3.38e+00;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1253 PMPOAPALWIEI 1264

OY 3 amdpptplwikt 14

RESULT 11  
ID YSCK\_YERS STANDARD; PRT; 209 AA.  
AC Q00927;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 37, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE YOP PROTEINS TRANSLOCATION PROTEIN K (LOW CALCIUM RESPONSE LOCUS DE PROTEIN KB).  
GN YSCK OR ICRKB.  
OS VERSINA PSEUDOTUBERCULOSIS.  
OS PLASMID PIBL.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
RN [1] VERSINA.  
RP SEQUENCE FROM N.A.  
RC STRAIN=YPIII;  
RX MEDLINE; 92250432.  
RA RIMPLAELNEN M.; FORSBERG A.; WOLF-WATZ H.;  
RT "A novel protein, IcrQ, involved in the low-calcium response of Yersinia pseudotuberculosis shows extensive homology to YopH.";  
RL J. BACTERIOL. 174:3355-3363(1992).  
CC -1- FUNCTION: BELONGS TO AN OPERON INVOLVED IN THE TRANSLLOCATION OF YOP PROTEINS ACROSS THE BACTERIAL MEMBRANES OR IN THE SPECIFIC CONTROL OF THIS FUNCTION.  
CC -1- INDUCTION: AT 37 DEGREES CELSIUS IN THE ABSENCE OF CALCIUM.  
CC -1- SIMILARITY: HIGH WITH Y ENTEROCOLITICA CORRESPONDING ORF.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi-sib.ch/announce/> or send an email to license@ebi-sib.ch).  
CC EMBL; M74011; G155560; -.  
DR PIR; B40049; B40049.  
KW PLASMID; VIRULENCE.  
SQ SEQUENCE 209 AA; 23998 MW; CDB29D0F CRC32;

Query Match 49.1%; Score 57; DB 1; Length 209;  
Best Local Similarity 53.8%; Pred. No. 5.12e+00;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
OY 1 lkdampptplwikt 13

RESULT 13  
ID RFE\_HAEIN STANDARD; PRT; 355 AA.  
AC P43341;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N-ACETYLGLUCOSAMINYLTRANSFERASE (EC 2.4.1.-)  
DE RFE OR H1716.  
GN HAEMOPHILUS INFLUENZAE.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;  
OC HAEMOPHILUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20;  
RX MEDLINE; 95550630.  
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
RA MCKENNEY K., SUTTON G., FINNUGH W., FIELDS C.A., GOCayne J.D.,  
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A., KELLEY J.M.,  
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEORGGEN N.S.M.,  
RA GINSER C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
RA VENTER J.C.;  
RT "Whole genome random sequencing and assembly of Haemophilus influenzae Rd.";  
RL SCIENCE 269:496-512(1995).  
CC -1- FUNCTION: MAY BE THE TONICAMICIN SENSITIVE TRANSFERASE THAT CATALYZES THE SYNTHESIS OF GLCNAc-PYROPHOSPHORYLUNDECAPRENOL (LIPID II), THE FIRST LIPID-LINKED INTERMEDIATE INVOLVED IN ECA SYNTHESIS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + UNDECAPRENYL PHOSPHATE = UMP + UNDECAPRENYL N-ACETYL-ALPHA-D-GLUCOSAMINYL PYROPHOSPHATE.  
CC -1- PATHWAY: INVOLVED IN THE SYNTHESIS OF ENTEROBACTERIAL COMMON ANTIGEN (ECA) AND REQUIRED FOR SYNTHESIS OF LIPOPOLYSACCHARIDE O-SIDE CHAINS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: SOME TO THE DOLICHYL-PHOSPHATE ALPHA-N-ACETYL-GLUCOSAMINYLTRANSFERASES OF YEAST AND MAMMALS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

DR EMBL; U32844; G1574571; -  
DR TIGR; H11716; -  
KW LIPOPOLYSACCHARIDE BIOSYNTHESIS; GLYCOSYLTRANSFERASE; TRANSFERASE;  
FT TRANSMEMBRAINE; 1 21 . POTENTIAL.  
FT TRANSMEM 39 59 . POTENTIAL.  
FT TRANSMEM 63 83 . POTENTIAL.  
FT TRANSMEM 100 120 . POTENTIAL.  
FT TRANSMEM 123 143 . POTENTIAL.  
FT TRANSMEM 151 171 . POTENTIAL.  
FT TRANSMEM 182 202 . POTENTIAL.  
FT TRANSMEM 208 228 . POTENTIAL.  
FT TRANSMEM 237 257 . POTENTIAL.  
FT TRANSMEM 289 309 . POTENTIAL.  
FT TRANSMEM 316 336 . POTENTIAL.  
SQ SEQUENCE 355 AA; 40070 MW; 4CALFF0A CRC32;

Query Match 49.1%; Score 57; DB 1; Length 355;

Best Local Similarity 50.0%; Pred. No. 5.12e+00; Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 236 PKNPVTALWI 245  
Qy 3 :||| :||| 12

RESULT 14 STANDARD; PRT; 192 AA.

ID Y4PG RHISN  
AC P5516;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 21.1 KD PROTEIN Y4PG/Y4SC.  
GN Y4PG AND Y4SC.  
OS RHIZOBIUM SP. (STRAIN NGR234).  
OC PLASMID SYM PGNR34A.  
OC BACTERIA; PROTEOBACTERIA; ALPPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
OC RHIZOBIACEAE; RHIZOBIUM.  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE; 97305956.  
RA FREIBERG C.A., FELLY R., BAIRACH A., BROUGHTON W.J., ROSENTHAL A.,  
RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
RL NATURE 387:394-401(1987);

-1- SIMILARITY: STRONG, TO ACETOBACTER XYLIUM IS1268 OREA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

EMBL; AE000090; G2182576; -  
DR HYPOTHETICAL PROTEIN; PLASMID; TRANSPOSABLE ELEMENT.  
SQ SEQUENCE 192 AA; 21094 MW; DB66859B CRC32;

Query Match 48.3%; Score 56; DB 1; Length 192;  
Best Local Similarity 55.7%; Pred. No. 7.73e+00; Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 166 SPPLWKE 174  
Qy 7 :||| :||| 15

RESULT 15 STANDARD; PRT; 328 AA.

ID SRGLCAEL  
AC P46570;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DR 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE SRG-1 PROTEIN.  
GN SRG-1 OR C18F10.4.

OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTIA; RHABDITIA; RHABDITIDA;  
RHABDITINA; RHABDITOIDEA; PELODERINAE; CAENORHABDITIS.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA LATREILLE P.;  
RL SUBMITTED (MAY-1994) TO EMBL/GENBANK/DDJB DATA BANKS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
CC -1- BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRG  
CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

EMBL; U00049; G485105; -  
DR WORMPP; C18F10.4; CB00711.

KW TRANSMEMBRANE; MULTIGENE FAMILY.  
FT TRANSMEM 25 45 . POTENTIAL.  
FT TRANSMEM 59 79 . POTENTIAL.  
FT TRANSMEM 118 138 . POTENTIAL.  
FT TRANSMEM 148 168 . POTENTIAL.  
FT TRANSMEM 197 217 . POTENTIAL.  
FT TRANSMEM 234 254 . POTENTIAL.  
FT TRANSMEM 298 318 . POTENTIAL.  
SQ SEQUENCE 328 AA; 38276 MW; 24F869EA CRC32;

Query Match 48.3%; Score 56; DB 1; Length 328;  
Best Local Similarity 50.0%; Pred. No. 7.73e+00; Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 152 IATMAISPCLW 163  
Qy 1 :||| :||| 12

Search completed: Wed Dec 29 14:06:44 1999  
Job time : 8 secs.







Best Local Similarity 50.0%; Pred. No. 6.29e+00; Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 10  
ENTRY | | 1:1:1  
Qy 1 lkandptpplw1 12

ALTERNATE\_NAMES major structural nucleoprotein - Machupo virus  
ORGANISM formal\_name Machupo virus  
#submission submitted to the EMBL DNA Library, October 1991  
#description sequence of the nucleocapsid gene of Machupo Virus: close relationship with another South American pathogenic arenavirus, Junin.

#accession S18042  
#molecule\_type genomic RNA  
#residues 1-564 #label GRI  
GENETICS  
#map\_position segment\_S  
CLASSIFICATION superfamily arenavirus major nucleoprotein  
KEYWORDS nucleocapsid; nucleoprotein  
SUMMARY #length 564 #molecular-weight 63299 #checksum 9818

Query Match 50.9%; Score 59; DB 1; Length 564;  
Best Local Similarity 46.7%; Pred. No. 6.29e+00;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 368 VKKIDPNTWLDIE 382  
Qy 1 lkandptpplw1 15

RESULT 11  
ENTRY B20807  
TIME complement C4B - human (fragment)  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 29-Aug-1997

ACCESSIONS  
REFERENCE  
#authors Belt, K.T.; Carroll, M.C.; Porter, R.R.  
#journal Cell (1984) 36:907-914  
#title The structural basis of the multiple forms of human complement component C4.

#cross-references MURD:81156544  
#accession B20807  
#molecule\_type mRNA  
#residues 1-184 #label BEL

GENETICS  
#gene GDB:C4B  
#map\_position GDB:119733; OMIM:120820  
CLASSIFICATION superfamily alpha-2-macroglobulin  
SUMMARY #length 184 #checksum 6928

Query Match 50.0%; Score 58; DB 2; Length 184;  
Best Local Similarity 50.0%; Pred. No. 9.22e+00; Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 137 PMPQAPALMWT 148  
Qy 3 vndptpplw1 14

RESULT 12  
ENTRY S65208  
TITLE probable membrane protein YPL189w - yeast (Saccharomyces cerevisiae)  
ALTERNATE\_NAMES hypothetical protein P2201  
ORGANISM formal\_name Saccharomyces cerevisiae  
#submission submitted to the Protein Sequence Database, May 1996  
#accession S65208  
#molecule\_type DNA  
#residues 1-609 #label RIE  
#cross-references EMBL:273545; NID:91370394; PID:e246916; PID:91370395;  
#authors Benes, V.; Rechmann, S.; Neutwich, U.; Voss, H.; Ansorge, W.  
#submission submitted to the Protein Sequence Database, May 1996  
#accession S65201  
#molecule\_type DNA  
#residues 177-609 #label BEN  
#cross-references EMBL:273545; MIPS:YPL189w  
#experimental\_source strain S288C (AB972)

GENETICS  
#map\_position 16L  
KEYWORDS transmembrane protein  
FEATURE 79-95  
#domain transmembrane #status predicted #label TM1  
136-152 #domain transmembrane #status predicted #label TM12  
164-180 #domain transmembrane #status predicted #label TM13  
201-217 #domain transmembrane #status predicted #label TM14  
328-344 #domain transmembrane #status predicted #label TM15  
376-392 #domain transmembrane #status predicted #label TM16  
406-422 #domain transmembrane #status predicted #label TM7  
498-514 #domain transmembrane #status predicted #label TM8  
534-550 #domain transmembrane #status predicted #label TM9  
578-594 #domain transmembrane #status predicted #label TM10  
SUMMARY #length 609 #molecular-weight 71288 #checksum 4342

Query Match 50.0%; Score 58; DB 2; Length 609;  
Best Local Similarity 54.5%; Pred. No. 9.22e+00;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 59 LSNNSPSPPLW 69  
Qy 1 :|:1:1:1  
1 lkandptpplw1 11

RESULT 13  
ENTRY C4HJ  
TITLE complement C4A precursor - human  
CONTAINS classical-complement-pathway C3/C5 convertase (EC 3.4.21.43)  
ORGANISM #formal\_name Homo sapiens #common\_name man  
#label C4a anaphylatoxin  
DATE 25-Feb-1985 #sequence\_revision 23-Aug-1996 #text\_change 24-Oct-1997

ACCESSIONS  
REFERENCE  
#authors Yu, C.Y.  
#journal J. Immunol. (1991) 146:1057-1066  
#title The complete exon-intron structure of a human complement component C4A gene. DNA sequences, polymorphism, and linkage to the 21-hydroxylase gene.

#cross-references MURD:91108039  
#accession 156095  
#status preliminary; translated from GB/EMBL/DDJB  
#molecule\_type DNA  
#residues 1-174 #label RES  
#cross-references GB:M59815; NID:9179672; PID:9179674



680-1446 #product complement C4 alpha chain #status predicted  
 #label ALR  
 #product C4a anaphylatoxin #status experimental #label  
 C4A  
 #region C4b-binding protein binding  
 #product C4d fragment #status experimental #label C4D  
 #product complement C4 gamma chain #status predicted  
 #label GAM  
 #cleavage\_site Arg-Ala (complement subcomponent C1s)  
 #status experimental  
 #binding\_site carbohydrate (Asn) (covalent) #status  
 experimental  
**SUMMARY**  
 #length 1744 #molecular-weight 192860 #checksum 9431  
 Query Match 50.0%; Score 58; DB 1; Length 1744;  
 Best Local Similarity 50.0%; Pred. No. 9.22e+00;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Date 17-Oct-1997  
 Db 1253 PMPCAPALIET 1264  
 Qy 3 amdpptplwikt 14  
**RESULT** 14  
**ENTRY** B4049 #type complete  
**TITLE** virC-region hypothetical protein ysck - Yersinia  
**ORGANISM** #formal\_name Yersinia enterocolitica  
**DATE** 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change  
**ACCESSIONS** 09-Sep-1997  
**REFERENCE** B4049  
 Michiels, T.; Vanooteghem, J.C.; Lambert de Rouvreit, C.;  
 China, B.; Gustin, A.; Boudry, P.; Cornelis, G.R.  
 J. Bacteriol. (1991) 173:4994-5009  
**TITLE** Analysis of virC, an operon involved in the secretion of Yop  
**#cross-references** MND:91317716  
**ACCESSION** B4049  
**#status** preliminary  
**#molecule\_type** DNA  
**#residues** 1-209 #label MIC  
**GENETICS** #cross-references GB:M74011; NID:9155549; PID:9155560  
**SUMMARY** #genome plasmid  
**LENGTH** 209 #molecular-weight 23998 #checksum 7033  
 Query Match 49.1%; Score 57; DB 2; Length 209;  
 Best Local Similarity 53.8%; Pred. No. 1.35e+01;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 Date 17-Oct-1997  
 Db 152 LAAMBPOPOWCK 164  
 Qy 1 lkamdpptplwikt 13  
**RESULT** 15  
**ENTRY** A64138 #type complete  
**TITLE** rfe protein - Haemophilus influenzae (strain Rd KW20)  
**ORGANISM** #formal\_name Haemophilus influenzae  
**DATE** 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change  
**ACCESSIONS** A64138  
**REFERENCE** A64008  
**#authors** Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;  
 Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.;  
 Dougherty, B.A.; Merrick, J.M.; McPenney, K.; Sutton, G.;  
 FitzHugh, W.; Fields, C.; Gooley, J.D.; Scott, J.;  
 Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,  
 J.F.; Phillips, C.A.; Spriggs, T.; Hedbom, E.; Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Sauder,  
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Funemann,  
 J.L.; Geoghegan, N.S.M.; Giehm, C.L.; McDonald, L.A.;  
 Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.  
**JOURNAL** Science (1995) 269:496-512  
**title** Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd  
**ACCESSION** A64138  
**#status** nucleic acid sequence not shown; translation not shown  
**#molecule\_type** DNA  
**#residues** 1-355 #label TIGR  
**#cross-references** MND:91530530  
 TIGR:HJ1716  
**GENETICS**  
**#gene** rfe  
**rfe** transmembrane protein  
**KEYWORDS** #length 355 #molecular-weight 40070 #checksum 678  
**SUMMARY**  
 Query Match 49.1%; Score 57; DB 2; Length 355;  
 Best Local Similarity 50.0%; Pred. No. 1.35e+01;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Date 17-Oct-1997  
 Db 236 PMKPYTAWI 245  
 Qy 3 amdpptplwikt 12  
**SEARCH**  
 Search completed: Wed Dec 29 14:06:19 1999  
 Job time : 16 secs.



OC SCIURGNATHI; MORIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 MEDLINE: 98359597.  
 YAMADA J., SUGA K., FURHATA T., KITAHARA M., WATANABE T.,  
 HOSOKAWA M., SATOH T., SUGA T.;  
 "cDNA cloning and genomic organization of peroxisome  
 proliferator-inducible long-chain acyl-CoA hydrolase from rat liver  
 cytosol";  
 BIOCHEM. BIOPHYS. RES. COMMUN. 248:608-612(1998).  
 EMBL: AJ010429; D|033502; -.  
 KW HYDROLASE.  
 NT NON-TER. 197 AA; 197  
 FT SEQUENCE 197 AA; 21593 MW; 47FBDE8C CRC32;  
 SQ

Query Match 56.0%; Score 65; DB 11; Length 197;  
 Best Local Similarity 54.5%; Pred. No. 2.14e-01;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 ikamadptpplw1 11

Db 122 IFRMEPERLW 132

RESULT 3 PRELIMINARY; PRT; 216 AA.  
 ID 032304; AC 032304;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE YIRD.  
 OS BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACTILLACEAE;  
 OC BACTILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 MEDLINE: 98015415.  
 MEDINA N., VANNIER F., ROCHE B., AUTRET S., LEVINE A., SEROR S.J.;  
 RT "Sequencing of regions downstream of adda (98 degrees) and citg (289  
 degrees) in *Bacillus subtilis*";  
 RL MICROBIOLOGY 143:3305-3308(1997).  
 DR EMBL: Z03941; E113561; -.  
 PRFM: PR00072; response/reg; 1.  
 DR PEAM: PR00486; trans-reg/C; 1.  
 SQ SEQUENCE 216 AA; 25130 MW; 26PA9BEB CRC32;

Query Match 55.2%; Score 64; DB 2; Length 216;  
 Best Local Similarity 63.6%; Pred. No. 3.37e-01;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Db 31 KKMPSPHLWI 41  
 Qy 2 kamadptpplw1 12

RESULT 4 PRELIMINARY; PRT; 225 AA.  
 ID 032192; AC 032192;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-NOV-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE YVQA PROTEIN.  
 GN BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACTILLACEAE;  
 OC BACTILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 MEDLINE: 98015415.  
 MEDINA N., VANNIER F., ROCHE B., AUTRET S., LEVINE A., SEROR S.J.;  
 RT "Sequencing of regions downstream of adda (98 degrees) and citg (289  
 degrees) in *Bacillus subtilis*";  
 RL MICROBIOLOGY 143:3305-3308(1997).  
 DR EMBL: Z03941; E113561; -.  
 PRFM: PR00072; response/reg; 1.  
 DR PEAM: PR00486; trans-reg/C; 1.  
 SQ SEQUENCE 225 AA; 26212 MW; FD1BD0D CRC32;

Query Match 55.2%; Score 64; DB 2; Length 225;  
 Best Local Similarity 63.6%; Pred. No. 3.37e-01;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Db 40 KKMPSPHLWI 50  
 Qy 2 kamadptpplw1 12

RESULT 5 PRELIMINARY; PRT; 836 AA.  
 ID P07359; AC P07359;  
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE TAII95.  
 OS MUS MUSCUS (MOUSE).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;  
 OC SCIURGNATHI; MORIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 MEDLINE: 9804033.  
 DR EMBL: Y09974; ED00061; -.  
 SQ SEQUENCE 836 AA; 92037 MW; A640D245 CRC32;

Query Match 55.2%; Score 64; DB 11; Length 836;

Best Local Similarity 50.0%; Pred. No. 3.37e-01; Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OS HOMO SAPIENS (HUMAN)

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;

OC CATARRHINI; HOMINIDAE; HOMO.

RN [1]

RA GUERRA M.; ROEDER R.G.R.;

RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.

DR EMBL; AF057694; G046652;

SQ SEQUENCE 1199 AA; 136985 MW; 98118DAB CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1199; Best Local Similarity 50.0%; Pred. No. 3.37e-01; Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

DB 615 LSAMDADSPLWIRID 631

QY | ||| :||| | : :

OC HOMO SAPIENS (HUMAN)

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;

OC CATARRHINI; HOMINIDAE; HOMO.

RN [1]

RP SEQUENCE FROM N.A.

RA MARTINEZ E.; GE H.; TAO Y.; YUAN C.-X.; ROEDER R.G.;

RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.

DR EMBL; AF040701; G2827282; -.

SQ SEQUENCE 1189 AA; 135833 MW; 180759F2 CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1189; Best Local Similarity 50.0%; Pred. No. 3.37e-01; Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

DB 606 LSAMDADSPLWIRID 621

QY | ||| :||| | : :

1 lkamdpptpl-wikte 15

---

RESULT 7 PRELIMINARY; PRT; 1199 AA.

ID 043487 AC 015070

AC 015070; PRELIMINARY; PRT; 1327 AA.

DT 01-JUN-1998 (TREMBLREL. 06, CREATED)

DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)

DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE KIAA0364.

GN HOMO SAPIENS (HUMAN)

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;

OC CATARRHINI; HOMINIDAE; HOMO.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE: 9734984.

RA NAGASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SEKI N., MIYAJIMA N., TANAKA A., KOBAYASHI H., NOMURA N., OHARA O.;

RT "Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."

RT DNA RES. 4:141-150(1997);

RL EMBL; AB002362; D1021660; -.

DR PF00047; 1g; 10.

SQ SEQUENCE 1327 AA; 147971 MW; 0A68A862 CRC32;

Query Match 55.2%; Score 64; DB 4; Length 1327; Best Local Similarity 58.3%; Pred. No. 3.37e-01; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 25 NDPOPELWIRSN 36

QY | ||| - ||| | : :

4 mdaptplwikte 15

---

RESULT 8 PRELIMINARY; PRT; 1199 AA.

ID 060668 AC 088550

AC 088550; PRELIMINARY; PRT; 97 AA.

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1998 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE VI, V2, C1, C2, C3 AND C4 GENES.

GN C4.

OC TOMATO YELLOW LEAF CURL VIRUS (TYLCV).

OC VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; BEGOMOVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=INDIAN TOMATO LEAF CURL VIRUS (ITMLCV);

RC MEDLINE: 95363124.

RA HONG Y.; HARRISON B.D.;

RT "Nucleotide sequences from tomato leaf curl viruses from different countries: evidence for three geographically separate branches in evolution of the coat protein of whitefly-transmitted geminiviruses"; J. GEN. VIROL. 76:2043-2049(1995).

RL TAII150.

DR EMBL; Z48182; G974214; -; SQ SEQUENCE 97 AA; 11026 MW; 62CE9CD2 CRC32;

Query Match 52.6%; Score 61; DB 14; Length 97;  
Best Local Similarity 46.7%; Pred. No. 1.29e+00; Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 42 LRAVQNSRPWAKTE 56  
QY |:!: : |:| |||  
1 ikandptppiwkt 15

---

RESULT 11 PRELIMINARY; PRT; 258 AA.  
ID 017984 AC 017984 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE R06B9\_2 PROTEIN.  
GN R06B9\_2.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA; RHABDITINAE; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1] SEQUENCE FROM N.A.  
RN BAYNES C.;  
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDJB DATA BANKS.  
RN [2] SEQUENCE FROM N.A.  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAINES C., BERKS M., COULSON A.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMLRRA J., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONHAMMER E., STADEN R., SULSTON J.,  
RA THIERR-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RA SWALDON N., SMITH A., SONHAMMER E., STADEN R., SULSTON J.,  
RA THIERR-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.",  
RA NATURE 368:32-38(1994).  
RN [2] SEQUENCE FROM N.A.  
RN STRAIN=BRISTOL N2;  
RA JOHNSON D., WAMSLEY P., KRAMER J., ELLIOTT G.;  
RT "The sequence of *C. elegans* cosmid TOBELL.";  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
RN [3] SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA WATERSTON R.;  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
DR EMBL; AF077546; G3319478; -;  
SQ SEQUENCE 354 AA; 40564 MW; 21878AC CRC32;

---

Query Match 50.9%; Score 59; DB 5; Length 354;  
Best Local Similarity 50.0%; Pred. No. 3.05e+00; Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 79 KHLDPAVPLWRKD 92  
QY |:| ||| |||  
2 kamdptppiwkt 15

---

RESULT 12 PRELIMINARY; PRT; 1220 AA;  
ID 098864 AC 098864 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
DE PATCHED PROTEIN.  
GN PTC1.

OS BRACHYDANTO RERIO (ZEBRAFISH) (ZEBRA DANTO).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPHYGII; NEOPTERYGII;  
OC TELOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;  
OC CYPRINIDAE; RASBORINAE; DANIO.  
RN [1] SEQUENCE FROM N.A.  
RA CONCORDETT J. P., LEWIS K. E., MOORE J. W., GOODRICH L. V., JOHNSON R. L.,  
SCOTT M. F., INGHAM P. W., GOODRICH L. V., JOHNSON R. L.,  
RA DEVELOPMENT 12:2835-2846(1996).  
RL EMBL; X98883; E25499; -;

---

RESULT 13 PRELIMINARY; PRT; 354 AA.  
ID 076537 AC 076537 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DE TOBELL.6 PROTEIN.  
GN TOBELL.6.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA; RHABDITINAE; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1] SEQUENCE FROM N.A.  
RN RC STRAIN=BRISTOL N2;  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAINES C., BERKS M., COULSON A.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMLRRA J., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONHAMMER E., STADEN R., SULSTON J.,  
RA THIERR-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.",  
RA NATURE 368:32-38(1994).  
RN [2] SEQUENCE FROM N.A.  
RN RC STRAIN=BRISTOL N2;  
RA JOHNSON D., WAMSLEY P., KRAMER J., ELLIOTT G.;  
RT "The sequence of *C. elegans* cosmid TOBELL.";  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
RN [3] SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA WATERSTON R.;  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
DR EMBL; AF077546; G3319478; -;  
SQ SEQUENCE 354 AA; 40564 MW; 21878AC CRC32;

---

Query Match 50.9%; Score 59; DB 5; Length 354;  
Best Local Similarity 50.0%; Pred. No. 3.05e+00; Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 45 LRHDPTPQFWF 56  
QY |:| ||| |||  
1 ikandptppiwkt 12

---

RESULT 14 PRELIMINARY; PRT; 453 AA.  
ID 055171 AC 055171 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL VERY-LONG-CHAIN ACYL-ICOA THIOESTERASE.  
OS RATTUS NORVEGicus (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
RN [1] SCIRROGANATHI; MURIDAE; MURINAE; RATTUS.

RP SEQUENCE FROM N.A.  
RC STRAIN="WISTAR;  
RX MEDLINE; 98109136.  
RA SVENSSON, ENGBERG, Aoyama, USUDA, ALEXSON, HASHIMOTO;  
RT "Molecular cloning and characterization of a mitochondrial peroxisome  
proliferator-induced acyl-CoA thioesterase from rat liver."  
RL BIOCHEM. J. 320:601-608(1998).  
DR EMBL; Y09333; E30716; \_  
SQ SEQUENCE 453 AA; 49701 MW; CFAEDB3 CRC32;

Query Match 50.9%; Score 59; DB 11; Length 453;  
Best Local Similarity 66.7%; Pred. No. 3.08e+00;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 3 amdpptpllw 11

RESULT 15

ID Q14033 PRELIMINARY; PRT; 100 AA.

AC 014033;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE COMPLEMENT COMPONENT C4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 8311735.

RA CARROLL, M.C., FORIER, R.R.;  
RT "Cloning of a human complement component C4 gene.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 80:264-267(1983).  
DR EMBL; V0502; E3882; -.  
DR PFAM; PF00207; A2M; 1.  
SQ SEQUENCE 100 AA; 10629 MW; 4589044E CRC32;

Query Match 50.0%; Score 58; DB 4; Length 100;  
Best Local Similarity 50.0%; Pred. No. 4.73e+00;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 59 PMPQAPALWIET 70  
Qy 3 :| :||| | 14

Search completed: wed Dec 29 14:07:20 1999  
Job time : 18 secs.



Matches	15; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	1 lkandptplwikte 15							
ID	W11575 standard; peptide; 15 AA.							
AC	W11575; US-05e387.							
DT	20-MAR-1997 (first entry)							
DE	N-terminal peptide from lethal toxin neutralising factor.							
KW	Lethal toxin neutralising factor; LTNF; opossum; bee toxin; scorpion toxin; plant toxin; bacterial toxin; venom; sting; snake bite.							
OS	bidelphis virginiana.							
PN	US5226209A							
PD	19-Nov-1996.							
PF	10-MAY-1993; 058387.							
PR	22-SEP-1994; US-310340.							
PA	(LIPPS/) LIPPS B V.							
PI	LIPPS BV LIPPS FW;							
DR	WPI; 97-011287/01.							
PT	treatment of victims of bee or scorpion stings or plant or bacterial toxins - by admin. of lethal toxin-neutralising factor or its N-terminal peptide							
PS	Claim 7; Column 11; 1lpp; English.							
CC	This sequence represents the peptide of the invention. It is a Lethal Toxin Neutralising Factor (LTNF) moiety from a 68 kDa anti-haemorrhagic protein derived from an opossum. The peptide can be used in a method for treating a victim of envenomation from a poisonous snake, preferably a poisonous snake from the family of Elapidae, Viperidae or sea snake.							
CC	It is useful for the treatment of snake bites, sepsis, allergies caused by the environment and treatment of bee or scorpion stings or toxicities caused by plant or bacterial toxins. The peptide can also be used in histamine reaction treatment. The peptide can be used in envenomation treatment for a variety of snakes without prior identification of the snake. Being short it can be synthetically prepared rather than the current production in horses, where some people can show hypersensitivity to horse proteins.							
SQ	Sequence 10 AA;							
Query Match	62.9%; Score 73; DB 30; Length 10;							
Best Local Similarity	100.0%; Pred. No. 8.51e-05;							
Matches	15; Conservative							
Db	1 lkandptplwikte 10							
ID	R56496 standard; Protein; 1189 AA.							
AC	R56496;							
DT	23-MAR-1995 (first entry)							
DE	TATA-binding Protein associated factor dTAF150; screening; diagnostic; therapeutic; gene transcription regulation.							
KW	dTAF150; screening; diagnostic; therapeutic; gene transcription regulation.							
OS	Drosophila.							
FH	Key location/qualifiers							
FT	misc_difference 923 /note= "Val or Ieu"							
FT	misc_difference 1106 /note= "Arg, Pro or His"							
FT	misc_difference 1172 /note= "STOP"							
FT	misc_difference 1176 /note= "STOP"							
FT	W09417087-A.							
PN	04-AUG-1994.							
PD	28-JAN-1994; U01114.							
PF	28-JAN-1993; US-013412.							
PR	30-JUN-1993; US-087119.							
PA	(REGC ) UNIV CALIFORNIA.							
PI	Comai L, Dynact BD, Hoey T, Ruppert S, Tanese N;							
PI	Tian R, Wang E, Weinzierl ROJ;							
DR	WPI; 94-264019/32.							
P-PSDB	970733.							
PT	TATA-binding protein associated protein factors - and corresponding nucleotide sequence and deriv. antibodies, useful in screening, diagnostics and therapeutics							
PT	Disclosure; Page 156; 180pp; English.							
CC	The TATA binding protein associated factor hTAF150 (including specific antibodies and fusion products) are used in drug screening, diagnostics and therapeutics. They are used in the development of specific biochemical assays for screening compounds that agonise or antagonise selected transcription factors involved in regulating gene expression associated with human pathology.							
SQ	Sequence 1189 AA;							
Query Match	62.1%; Score 72; DB 11; Length 1189;							
Best Local Similarity	53.3%; Pred. No. 3.58e+00;							
Matches	8; Conservative							
Db	589 lsamdsppvlwrld 603							
QY	1 lkandptplwikte 15							

RESULT 3  
ID W53843 standard; peptide; 10 AA.  
AC W53843;  
DT 08-JUL-1998 (first entry)  
N-terminus of opossum LTNF.  
KW LTNF; lethal toxin neutralising factor; opossum; envenomation; therapy; anti-haemorrhagic protein; Elapidae; Viperidae; sea snake; snake bite; sepsis; allergy; bee sting; scorpion sting; plant toxin; bacterial toxin; histamine reaction treatment.  
OS Didelphis virginiana.  
PN US544444-A.  
PD 28-APR-1998.  
PF 03-JUN-1996; 657163.  
PR 03-JUN-1996; US-657163.  
03-JUN-1993; US-05e387.  
22-SEP-1994; US-310340.  
(LIPPS/) LIPPS B V.  
LIPPS BV LIPPS FW.  
DR WPI; 98-271108/24.  
PT lethal toxin Neutralising Factor peptide from opossum - can neutralise venom(s) from all major families of poisonous snakes

**RESULT 5**  
 ID W25029 standard; Protein; 1213 AA.  
 AC W25029;  
 DT 08-OCT-1997 (first entry)  
 DE TATA binding protein associated factor, dTAFII150.  
 KW RNA polymerase transcription; TATA-binding protein;  
 KW initiation.  
 OS Drosophila sp.  
 PN US5637686A.  
 PD 10-JUN-1997.  
 PP 28-JAN-1993; US-013412.  
 PR 30-JUN-1993; US-087119.  
 PR 28-JAN-1994; US-18882.  
 PR 09-MAY-1996; US-646715.  
 PA (REGC ) UNIV CALIFORNIA.  
 PT Comai L, Dynlach BD, Hoey T, Ruppert S, Tanese N;  
 PT TJian R, Wang E, Weinzierl ROJ;  
 DR WPI; 97-319113-29.  
 DR N-PSDB; T#9605.

**PT** Nucleic acids encoding human TATA-binding protein associated factor (TAF) peptide(s) for production of recombinant peptide(s), used for modulating transcription of TAFs

**PS** Example 1: Column 131-138: 86pp; English.

**CC** W25029 represents TATA binding protein associated factor (TAF) polypeptide, dTAFII150 (mol. weight 150kD). TAF peptides derived from dTAFII30 alpha, dTAFII30 beta, dTAFII40, dTAFII60, dTAFII80, dTAFII10, dTAFII150, and dTAFII250, their human equivalents and nucleic acids encoding them, are used to modulate transcription, including transcription initiation. TAFs are nuclear proteins involved in RNA polymerase I, II and III transcription. The peptides act by binding to a different TAF, an activator, or TBP (TATA-binding protein) compound, typically a protein like TBP or another TAF, an activator, or DNA.

**SQ** Sequence 1213 AA;

Query Match 62.1%; Score 72; DB 23; Length 1213;  
 Best Local Similarity 53.3%; Pred. No. 3.58e+00;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 620 lsamadspvwlwrd 634  
 Qy |||| :||| :|||:  
 1 lkamadptpplwkte 15

**RESULT 6**  
 ID W06086 standard; protein; 1213 AA.  
 AC W06086;  
 DT 27-JAN-1997 (first entry)  
 DE Drosophila TATA-binding protein associated factor dTAFII150 Protein.  
 KW Drosophila; TATA-binding protein; TBP associated factor; TFIID;  
 KW RNA polymerase II; transcription; messenger RNA; nuclear fraction;  
 KW holoenzyme; lambda-gt1; expression library.  
 OS Drosophila melanogaster.  
 PN US553410-A.  
 PD 09-JUL-1996.  
 PP 28-JAN-1993; 013412.  
 PR 28-JAN-1993; US-013412.  
 PR 30-JUN-1993; US-087119.  
 PR 28-JAN-1994; US-18882.  
 PA (REGC ) UNIV CALIFORNIA.  
 PT Comai L, Dynlach BD, Hoey T, Ruppert S, Tanese N;  
 PT TJian R, Wang E, Weinzierl ROJ;  
 DR WPI; 96-33345/33.  
 Screen for cpds. that bind human TATA-binding protein associated factor - by testing ability to bind to polypeptide fragments of the factor, useful as (ant)agonists of transcription factors involved in disease.

**Examples:** Column 123-132; 86pp; English.

This is the amino acid sequence of the drosophila TATA-binding protein

**RESULT 7**  
 ID W29670 standard; Protein; 222 AA.  
 AC W29670;  
 DT 09-NOV-1998 (first entry)  
 DE Homo sapiens clone AM42\_3 secreted protein.  
 KW Clone; secreted protein.  
 OS Homo sapiens.  
 PT Key Location/Qualifiers  
 FT Peptide 2.14  
 FT /note= "signal peptide"  
 PN WO9832853-A2.  
 PD 30-JUL-1998.  
 PR 23-JAN-1998; U01396.  
 PR 24-JAN-1997; US-788789.  
 PA (GEMY ) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lawallie ER, McCoy JM, Merberg D,  
 PI Racine LA, Spaulding V, Treacy M;  
 DR WPI; 98-427949/36.  
 N-PSDB; V40540.  
 PT New isolated polynucleotide(s) and secreted proteins - isolated from human foetal kidney, adult brain, adult salivary gland, foetal brain and adult testes cDNA libraries  
 PR Claim 15; Page 65-66; 109pp; English.  
 PS The sequence is that of a secreted protein. Such a protein can have biological activities, e.g. nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating or suppressing activity, haemotopesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, tumour inhibition activity, and other activities.  
 CC Sequence 222 AA;

Query Match 51.7%; Score 60; DB 34; Length 222;  
 Best Local Similarity 60.0%; Pred. No. 5.27e+11;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 135 pmadpsypvi 144  
 Qy 3 :|||: |||:  
 3 andptpplw 12

**KW** Tobacco leaf curling virus gene; **TLCV**; promoter; C4 protein.  
**OS** Tobacco leaf curling virus.

**PN** JI007982-A.  
**PD** 11-MAR-1998.  
**PF** 30-AUG-1996; JP-230394.  
**PR** (NORQ ) NORINSUIRANSHO KYUSHU NOYO SHIKENSTO.  
**DR** WPI; 98-233630/21.  
**PT** Tobacco leaf curling virus gene - useful for inserting into vectors  
for expression in, e.g. tomato plants

**Example 4; Fig 5; 9pp; Japanese.**

This sequence represents the C4 protein encoded by the tobacco leaf curling virus (TLCV) gene of the invention. TLCV gene or its promoter can be inserted into a vector for expression in plants, e.g. tobacco and tomato. This sequence is believed to be encoded by the TLCV gene shown in V29761.

Revised entry submitted to correct the Patent Number Field

**SQ** Sequence 97 AA;

**Query Match** 48.3%; Score 56; DB 32; Length 97;  
**Best Local Similarity** 53.3%; Pred. No. 1.25e+02;  
**Matches** 8; **Conservative** 3; **Mismatches** 4; **Indels** 0; **Gaps** 0;

**Db** 42 lkarnqmsptwktte 56

**Qy** ||| : ::||| :|||  
1 lkamdpptplwiktte 15

**RESULT** 9

**ID** W58530 standard; Protein; 311 AA.

**AC** W58530;  
**DT** 01-SEP-1998 (first entry)

**DE** Peroxisomal thioesterase.

**KW** Human; peroxisomal thioesterase; **PxTE**; diagnosis; inflammatory disease; fatty acid metabolism disorder; cancer; drug screening.

**OS** Homo sapiens.

**US577653-A.**  
07-JUL-1998.  
11-JUN-1997; 872784.  
11-JUN-1997; US-872784.

**PA** (INCY-) INCITE PHARM INC.

**PI** Corley NC, Hillman JL, Shah P;  
WRI, 98-398039/34.

**DR** WPI; 98-398039/34.

**PT** New nucleic acid encoding human peroxisomal thioesterase - useful for diagnosis, treatment and prevention of fatty acid metabolism disorders, cancer and inflammatory disease, and in drug screening

**Claim 9; Fig 1; 30pp; English.**

The present sequence represents human peroxisomal thioesterase (PxTE). The cDNA encoding PxTE was isolated from a human BRAIN/NG clone 2150905 cDNA library. PxTE and its agonists are used in treatment, diagnosis and prevention of disorders of fatty acid metabolism e.g. neuronal disorders, Zellweger or Down's syndrome, Refsum's disease, epilepsy, Alzheimer's disease, depression, tardive dyskinesia, multiple sclerosis, Parkinson's disease, schizophrenia and Tourette's syndrome. Fragments of PxTE DNA can also be used for diagnosis and monitoring (in hybridisation or amplification methods), including determining a predisposition to disease, for gene mapping and to detect PxTE-encoding nucleic acid, particularly after amplification. PxTE can be used to raise antibodies (Ab) used therapeutically, in competitive drug screens, as immunoassay reagents for diagnosis and monitoring of diseases and for purification of natural PxTE, and to screen for specific binding agents antagonists e.g. specific Ab, antisense sequences or ribozymes which are used in treatment of cancer and inflammatory diseases (e.g. adult respiratory distress syndrome, asthma, diabetes mellitus, microbial and other infections). Administration of proteins, agonist or antagonists, optionally combined with conventional pharmaceuticals providing a synergistic effect, is e.g. orally, by injection or topically.

**SQ** Sequence 311 AA;

**Query Match** 47.4%; Score 55; DB 32; Length 311;  
**Best Local Similarity** 28.6%; Pred. No. 1.55e+02;  
**Matches** 4; **Conservative** 5; **Mismatches** 5; **Indels** 0; **Gaps** 0;

**RESULT** 10

**ID** W20095 standard; Protein; 446 AA.

**AC** W20095;  
**DT** 08-JUL-1997 (first entry)

**DE** H. pylori cytoplasmic protein 10677187-aa.

**KW** Cytoplasmic; vaccine; prevention; treatment; infection; identification; binding compound; bacterium; life cycle; activator; bacteria; inhibitor; quodent ulcer disease; chronic gastritis; diagnosis; envelope.

**OS** Helicobacter pylori.

**PN** WO964893-A1.  
06-DEC-1996.  
07-JUN-1996; U00122.  
01-APR-1996; US-630405.

**PA** (ASTR ) ASTRA AB.

**PI** Berglindh OT, Smith D, Nellagaard BL;

**DR** WPI; 97-052306/05.

**PT** Helicobacter pylori nucleic acid sequences and related polypeptides(s) - useful for vaccines to treat or prevent H. pylori infection, and to detect Helicobacter

**PS** Claim 61; Pages 329-330; 1B1PP; English.

This sequence is a H. pylori cytoplasmic protein involved in genome replication, translocation and repair.

**CC** The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors.

**CC** The genomic sequence of H. pylori (ATCC 5569) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.

**SQ** Sequence 446 AA;

**Query Match** 47.4%; Score 55; DB 21; Length 446;  
**Best Local Similarity** 50.0%; Pred. No. 1.55e+02;  
**Matches** 5; **Conservative** 3; **Mismatches** 2; **Indels** 0; **Gaps** 0;

**Db** 403 akepppnlw 412

**Qy** ||| : ::||| :|||:  
3 andptplwli 12

**RESULT** 11

**ID** R71375 standard; Protein; 551 AA.

**AC** R71375;  
**DT** 21-NOV-1995 (first entry)

**DE** Human cystathione beta synthase.

**KW** Cystathione; beta synthase; human; homocystinuria.

**OS** Homo sapiens.

**PN** WO950714-A.  
23-MAR-1995.  
12-SEP-1994; U10203.

**PR** 13-SEP-1993; US-120960.

**PA** (COLS ) UNIV COLORADO.

**PI** Kraus JP;  
WPI; 95-131185/17.

**DR** WPI; 95-131185/17.

**N-PSDB**; 087430.

**PT** purified DNA encoding human cystathione beta synthase, used for producing human cystathione beta synthase, used for treating homocystinuria

**Claim 7; Fig 1a-c; 45pp; English.**

CC The amino acid sequence shown in R11376 is the human cystathione  
beta-synthase (CBS). Human CBS can be used to treat patients with  
homocystinuria and the cDNA sequence, Q8Y430, from which it is  
derived is also useful for screening CBS deficient patients for  
sequence 551 AA;

CC

SQ

Query Match 47.4%; Score 55; DB 13; Length 551;  
Best Local Similarity 33.3%; Pred. No. 1.55e+02; Mismatches 3; Indels 0; Gaps 0;  
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Db 36 keapeplwird 47  
Qy ::|||: :  
4 mdptpplwikte 15

RESULT 12

ID W21020 standard; protein; 594 AA.

AC W21020;

DT 22-JUL-1997 (first entry)

DE H. pylori cytoplasmic protein, hpse1540orf18.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification; binding compound; bacterium; life cycle; activator; bacteria; inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis; envelope; genome; replication; transcription; recombination; repair.

KW Helicobacter pylori.

OS W09640893-A1.

PN 19-DEC-1996.

PR 07-JUN-1996; U09122.

PR 01-APR-1996; US-630405.

PR (ASTR ) ASTRA AB

PI Berglindh OI, Smith D, Mellgaard BL;

WPI: 97-052306/05.

DR N-PSDB; T65273.

PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) - useful for vaccines to treat or prevent H. pylori infection, and to detect Helicobacter

PS Claim 61; Pages 1410-1411; 1481pp; English.

CC The present sequence is a Helicobacter pylori cytoplasmic protein involved in genomic replication, transcription, recombination and repair. The protein may be used in a vaccine to prevent or treat H. pylori.

CC infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 100 nucleotides, and the predicted coding regions defined by computer evaluation. To identify

CC likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.

CC Sequence 594 AA;

SQ

Query Match 47.4%; Score 55; DB 22; Length 594;

Best Local Similarity 50.0%; Pred. No. 1.55e+02; Mismatches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 451 akepppnlw 460  
Qy ::|||: :  
3 mdptpplw 12

RESULT 13

ID R04197 standard; protein; 769 AA.

AC R04197;

DT 12-JUN-1990 (first entry)

DE Env gene of simian immunodeficiency virus.

KW SIV; simian immunodeficiency virus; AIDS; HIV; vaccine; ss.

OS SIV. J01289485-A.

PD 21-NOV-1989.

PF 16-JUN-1988; 1149023.

PR 16-TUN-1988; JP-119023.

PA (TOFU) Tca Nenryo Kogyo KK.

DR WPI; 90-005176501.

DR N-PSDB; 002829.

PT DNA complementary to RNA of simian immuno-deficiency virus - used in vaccines and for diagnosis of AIDS.

PT Disclosure: FIG 4; 6pp; Japanese.

CC Env gene derived from RNA of SIV from which vector plasmid PSAH121

may be constructed and incorporated into an E.coli vector.

CC Useful in developing a vaccine against and in diagnosis of AIDS.

Query Match 47.4%; Score 55; DB 2; Length 769;  
Best Local Similarity 50.0%; Pred. No. 1.55e+02; Mismatches 6; Conservative 2; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Db 43 mtpttslwattn 54  
Qy ::|||: :  
4 mdptpplwikte 15

RESULT 14

ID W22986 standard; Protein; 233 AA.

AC W22986;

DT 08-OCT-1997 (first entry)

DE Human serine protease 60 (SP60).

KW Human; colon carcinoma; COLO 201; cell line; serine protease; SP60; screening; inhibitor; treatment; disease.

OS Homo sapiens.

PN J0919790-A.

PR 10-JUN-1997.

PR 24-JUL-1997; 212195.

PR 29-SEP-1995; JP-275105.

PR (SUNRI ) SUNTORY LTD.

DR WPI; 97-357902/33.

DR N-PSDB; T79127.

PT Human colon carcinoma derived serine protease(s) SP59, SP60 and SP67 - useful to screen for specific inhibitors, e.g. to search for, or study agent for treatment of various diseases.

PT Claim 1; Pages 11-12; 14pp; Japanese.

CC The present sequence is the human colon carcinoma COLO 201 cell line derived serine protease 60 (SP60), which can be used to screen for specific inhibitors, e.g. to search for, or study an agent for the treatment of various diseases.

CC Sequence 233 AA;

SQ

Query Match 46.6%; Score 54; DB 23; Length 233;

Best Local Similarity 55.6%; Pred. No. 1.91e+02; Mismatches 5; Conservative 2; Indels 0; Gaps 0;

Db 114 ltpatplwl 122  
Qy ::|||: :  
4 mdptpplw 12

RESULT 15

ID W56695 standard; Protein; 396 AA.

AC W56695;

DT 24-JUN-1998 (first entry)

DE Tetracycline resistance sequence contained in plasmid PRZTLL.

KW Tn5 transposase; modified; enzyme; in vitro transposition; mutant; target; marker; transposon 5; Plasmid PRZTLL; tetracycline resistance.

OS Synthetic.

OS Escherichia coli.

PN W09810077-A1.

PD 12-MAR-1998.

PF 09-SEP-1997; UI5941.

PR 02-MAY-1997; US-850880.

PR 09-SEP-1996; US-814877.

PA (WISC ) WISCONSIN ALUMNI RES FOUND.

PI Goryshin IV, Reznikoff WS, Zhou H;

DR WPI; 98-193627/17.  
 DR N-PSDB; V28398.  
 PT Modified Tn5 transposase construct used in novel system for in vitro  
 PT transposition - used to, e.g. create absolute defective mutants,  
 PT provide selective markers and to facilitate insertion of specialised  
 PT DNA sequences into target DNA.  
 PS Disclosure: Pages 35-46; 73pp; English.  
 CC This tetracycline resistance sequence is contained in the plasmid  
 CC pZP7UL which is used to demonstrate in vitro transposition of a  
 CC transposable element located between a pair of Tn5 (transposon 5)  
 CC outside end (OE) termini. The invention provides a genetic construct  
 CC that contains a nucleotide sequence encoding a modified Tn5 transposase  
 CC enzyme that has both greater avidity for Tn5 OE repeats and is less  
 CC likely to assume an inactive multimeric form than a wild type Tn5  
 CC transposase and a transposable DNA sequence flanked at its 5' and 3' ends  
 CC by an 18 or 19 base pair flanking DNA sequence comprising nucleotide A  
 CC at position 10, T at 11 and A at 12. The modified Tn5 transposase and  
 CC the transposable DNA which is a DNA donor molecule are used in a system  
 CC for in vitro transposition. The system and method can be used to create  
 CC absolute defective mutants, to provide selective markers to target DNA,  
 CC to provide portable regions of homology to a target DNA, to facilitate  
 CC insertion of specialised DNA sequences into target DNA, to provide primer  
 CC binding sites or tags for DNA sequencing, to facilitate production of  
 CC genetic fusion for gene expression studies and protein domain mapping, as  
 CC well as to bring together other desired combinations of DNA sequences  
 CC (combinatorial genetics). The modified Tn5 transposase facilitates in  
 CC vitro transposition reaction rates of at least about 100-fold higher  
 CC than can be achieved using wild type transposase (as measure in vivo).  
 CC In vitro transposition using this system can also use donor DNA and  
 CC target DNA that is circular or linear. The system also requires no  
 CC outside high energy source and no other protein other than the modified  
 CC transposase.  
 SQ Sequence 396 AA;

Query Match 46.6%; Score 54; DB 30; Length 396;  
 Best Local Similarity 66.7%; Pred. No. 1.91e+02;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Db 90 mattpvlwi 98  
 Qy 4 mdptpplwi 12

Search completed: Wed Dec 29 14:05:47 1999  
 Job time : 29 secs.